

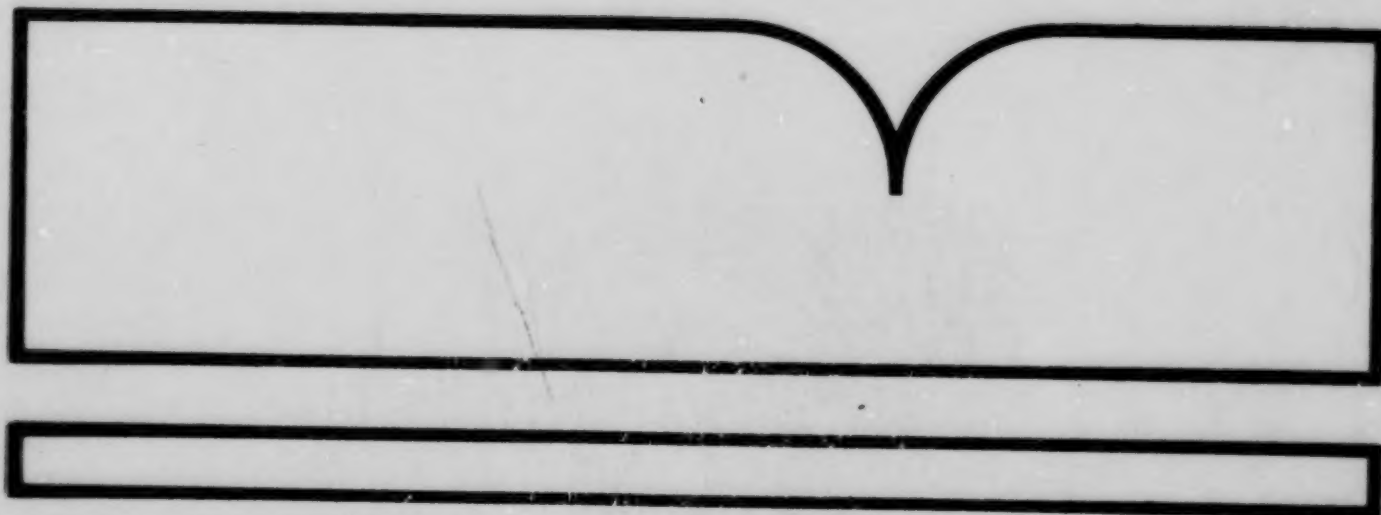
Economic Modelling of Water Supply
An Econometric Analysis of the
Multiproduct Firm

Cincinnati Univ., OH

Prepared for

Environmental Protection Agency, Cincinnati, OH

Mar 85



U.S. Department of Commerce
National Technical Information Service

NTIS

EPA/600/2-85/019
March 1985

ECONOMIC MODELLING OF WATER SUPPLY:
AN ECONOMETRIC ANALYSIS OF THE MULTIPRODUCT FIRM

by

Hak Youn Kim
Department of Economics
College of Arts and Sciences
University of Cincinnati
Cincinnati, OH 45221

Cooperative Agreement No. CR808102020
Project Period January 1981 - September 1983

Project Officer
Robert M. Clark
Drinking Water Research Division
Water Engineering Research Laboratory
U.S. Environmental Protection Agency
Cincinnati, OH 45268

WATER ENGINEERING RESEARCH LABORATORY
OFFICE OF RESEARCH AND DEVELOPMENT
U.S. ENVIRONMENTAL PROTECTION AGENCY
CINCINNATI, OH 45268

TECHNICAL REPORT DATA (Please read Instructions on the reverse before completing)		
1. REPORT NO. EPA/600/2-85/019	2.	3. RECIPIENT'S ACCESSION NO. PB85 176899/AS
4. TITLE AND SUBTITLE Economic Modelling of Water Supply: An Econometric Analysis of the Multiproduct Firm	5. REPORT DATE March 1985	
	6. PERFORMING ORGANIZATION CODE	
7. AUTHOR(S) Hak Youn Kim	8. PERFORMING ORGANIZATION REPORT NO.	
9. PERFORMING ORGANIZATION NAME AND ADDRESS Department of Economics College of Arts and Sciences University of Cincinnati Cincinnati, OH 45221	10. PROGRAM ELEMENT NO.	
	11. CONTRACT/GRANT NO. CR808102-02	
12. SPONSORING AGENCY NAME AND ADDRESS Municipal Environmental Research Laboratory - Cin, OH Office of Research and Development U.S. Environmental Protection Agency Cincinnati, Ohio 45268	13. TYPE OF REPORT AND PERIOD COVERED	
	14. SPONSORING AGENCY CODE EPA/600/14	
15. SUPPLEMENTARY NOTES Project Officer: Robert M. Clark 513/684-7228		
16. ABSTRACT Research was conducted to develop a comprehensive economic model that could use the neoclassical theory of the multiproduct firm to analyze the production structure of water supply. The project attempts to meet the need for in-depth analysis of the cost and economic structure of water supply. The cost structure of the water supply industry was estimated using the recently refined translog function. Issues addressed included restrictive specifications for water supply technology, input demand function for water supply, scale economies, marginal costs and output supply, and pricing of water and the presence of cross subsidization among residential and nonresidential users. Findings of this research represents the first attempt to apply the translog production function in the water supply area. No comprehensive analysis has been conducted before using the econometric methodology adopted in this study.		
17. KEY WORDS AND DOCUMENT ANALYSIS		
a. DESCRIPTORS	b. IDENTIFIERS/OPEN ENDED TERMS	c. COSATI Field/Group
18. DISTRIBUTION STATEMENT RELEASE TO PUBLIC	19. SECURITY CLASS (This Report) UNCLASSIFIED	21. NO. OF PAGES 228
	20. SECURITY CLASS (This page) UNCLASSIFIED	22. PRICE

DISCLAIMER

The information in this document has been funded wholly or in part by the United States Environmental Protection Agency under Cooperative Agreement No. CR808102020 to the University of Cincinnati. It has been subject to the Agency's peer and administrative review, and it has been approved for publication as an EPA document. Mention of trade names or commercial products does not constitute endorsement or recommendation of use.

TABLE OF CONTENTS

DISCLAIMER	ii
FOREWORD	iii
ABSTRACT	iv
ACKNOWLEDGMENT.....	ix
LIST OF TABLES.....	x
LIST OF FIGURES.....	xi
CHAPTER 1 INTRODUCTION	1
1.1 Nature of the Problem	1
1.2 Purpose of the Research	2
1.3 Methodology of the Research	5
1.4 Plan of the Research	6
CHAPTER 2 WATER SUPPLY INDUSTRY	8
2.1 Characteristics of Water Supply Systems	9
2.2 Problems of Water Supply and Water Supply Systems ..	16
2.3 Relevant Literature Review	23
2.4 Problems of Existing Studies	28
CHAPTER 3 FORMULATION OF THE MODEL	32
3.1 Characterization of the Cost Function for Water Supply	33
3.2 Translog Specification of the Model	49
3.3 Tests of the Underlying Production Structure of Water Supply	57
3.4 Structure of Input Demand	68
3.5 Scale Economies	74
3.6 Relation Between Operating Variables and Costs	78
3.7 Marginal Costs and Output Supply	80
3.8 Second Best Pricing Rule	86
Appendix: Derivation of Parameter Restrictions for Homo- geneity in Input Prices and Outputs	93
CHAPTER 4 EMPIRICAL METHODOLOGY	96
4.1 Econometric Specification and Estimation Procedure .	97
4.2 Source and Construction of Data	108
4.3 Analysis of the Sample Data	113

CHAPTER 5	EMPIRICAL FINDINGS	116
5.1	Estimated Cost Functions and Their Properties	117
5.2	Structure of Input Demand	124
5.3	Scale Economies	142
5.4	Relationship Between Operating Variables and Costs .	158
5.5	Marginal Costs and Output Supply	164
5.6	Pricing of Water and Cross-Subsidization Among Customer Classes	172
CHAPTER 6	SUMMARY AND CONCLUSIONS	188
6.1	Major Findings of the Research	189
6.2	Suggestions for Future Research	192
BIBLIOGRAPHY		196

LIST OF TABLES

Table 2.1	Central Water Systems and Population Served in the United States	13
Table 2.2	Operating Characteristics of Water Systems	17
Table 2.3	Finances of Government-Operated Water Utilities By Type of Government - 1976/1977	18
Table 4.1	Sample Statistics	114
Table 5.1	Parameter Estimates for the General Translog Multi-product Cost Function	118
Table 5.2	Test Statistics for Restricted Models	122
Table 5.3	Elasticities of Substitution Evaluated at Sample Means	126
Table 5.4	Own and Cross Price Elasticities of Input Demand Evaluated at Sample Means	129
Table 5.5	Elasticities of Input Demands With Respect to Changes in Demand for Water (Residential and Non-Residential), Capacity Utilization, and Service Distance Evaluated at Sample Means	137
Table 5.6	Overall Scale Economies for Selected Utilities	151
Table 5.7	Relationship Among System Size, Service Distance and Scale Economies	157
Table 5.8	Cost Elasticity of Service Distances and Corresponding Marginal Cost for Selected Utilities	163
Table 5.9	Marginal Cost Elasticities With Respect to Outputs, Input Prices, and Operating Variables Evaluated at Sample Means	168
Table 5.10	Actual Price-Marginal Cost Margins and Markup	178
Table 5.11	Second Best Prices Under Alternative Price Elasticities of Water Demand	182
Table 5.12	Relation of Actual to Second Best Prices	183

LIST OF FIGURES

Figure 1	A Typical Water Supply System	10
Figure 2	Tests of Underlying Production Structure Using Translog Cost Functions	58
Figure 3	Transformation Frontier	173

ACKNOWLEDGMENTS

The completion of this research would have been impossible without the assistance and impetus provided by the many people. In particular, Dr. Robert Clark (EPA project officer) was most instrumental in finalization of this work with his generous support and encouragement. There is no adequate way to express my gratitude to Bob.

I extend my sincere appreciation to Dr. Wolfgang Mayer and Dr. Haynes Goddard of the University of Cincinnati for questions and comments at the final stage of this research.

I owe my special thanks to Dick Stevie, John Ireland, Jim Goodrich, and other DWRD members of the EPA. Their time and efforts and, even more, friendship were a positive externality to the preparation of this research. Some of the data for this work was obtained with the cooperation of Jim Gillean of Act Systems, Inc., who is to thanked as well.

previous research, the water supply industry experiences remarkable economies associated with the treatment of water but suffers from diseconomies of increasing distribution network size. However, the diseconomies of distribution do not offset economies of treatment.

Marginal costs of residential and nonresidential usage of water are much lower than the actual prices charged. Marginal costs of output vary not only with customer classes but over various distances as well. In addition, the transformation function of water supply shows that there is a potential gain achievable from complete specialization in the production of either residential or nonresidential water supply.

Since water utilities enjoy marked economies of scale, this suggests the infeasibility of marginal cost pricing. The second-best (Ramsey optimal) prices are calculated for residential and nonresidential water usage. Surprisingly, the existing price structure is close to the second-best pricing optimum. Furthermore, this study finds no cross-subsidization among residential and nonresidential users.

This report was submitted in fulfillment of Cooperative Agreement No. CR808102020 by the University of Cincinnati under the sponsorship of the U.S. Environmental Protection Agency. This report covers the period January 1981 to September 1983, and work was completed as of August 1983.

is utilized which facilitates application of the translog multiproduct cost function. An unusual feature of the model developed via this research is that a few operating characteristics of the firm as measured by capacity utilization and service distance are included. The data employed come from a cross section of 1973 U.S. water utilities. The iterative Zellner efficient estimation procedure is utilized to obtain maximum likelihood estimators.

Estimation presumes nonhomotheticity among outputs (residential and nonresidential water supply), input prices, and operating variables in the production of delivered water. The evidence suggests that existing specifications placing restrictions on the translog multiproduct cost function are inappropriate. The tests of homogeneity of outputs, input-output separability, nonjointness, and a Cobb-Douglas form produce results that are significantly different from the unconstrained translog model.

Estimates of the elasticity of substitution indicate that capital is a substitute for energy and labor and that no strong substitution exists between energy and labor. Estimates of the output elasticity of input demands suggest that the production of water is an energy-intensive process. Further results show that efforts to analyze input demands without considering capacity utilization and service distance would lead to a serious bias in the estimation of water supply input requirements.

Water utilities are found to experience substantial economies of scale for both residential and nonresidential water supply customers, a finding which supports previous investigations. As conjectured in the

ABSTRACT

ECONOMIC MODELLING OF WATER SUPPLY: AN ECONOMETRIC ANALYSIS OF THE MULTIPRODUCT FIRM

By

Hak Youn Kim

Despite the fact that many firms market a multiplicity of products, studies on the multiproduct nature of the firm are rather limited. The purpose of this research is to develop a comprehensive economic model suitable for applying the neoclassical theory of the multiproduct firm to the analysis of the production structure of water supply. The following issues are addressed in the model:

- (1) restrictiveness of existing specifications of water supply technology in view of the neoclassical theory of production;
- (2) investigation of input substitution and utilization in the production process;
- (3) existence of scale economies;
- (4) examination of marginal costs and the transformation function of water supply;
- and (5) applicability of second-best pricing and presence of cross-subsidization among customer classes.

The water supply utility is viewed as a regulated multiproduct firm providing distinct services to different customer classes such as residential and nonresidential users. Capital, labor, and energy are considered as the major inputs into the production process for water delivered to meet customer demands. The principle of duality

FOREWORD

The U.S. Environmental Protection Agency is charged by Congress with protecting the Nation's land, air, and water systems. Under a mandate of national environmental laws, the agency strives to formulate and implement actions leading to a compatible balance between human activities and the ability of natural systems to support and nurture life. The Clean Water Act, the Safe Drinking Water Act, and the Toxics Substances Control Act are three of the major congressional laws that provide the framework for restoring and maintaining the integrity of our Nation's water, for preserving and enhancing the water we drink, and for protecting the environment from toxic substances. These laws direct the EPA to perform research to define our environmental problems, measure the impacts, and search for solutions.

The Water Engineering Research Laboratory is that component of EPA's Research and Development program concerned with preventing, treating, and managing municipal and industrial wastewater discharges; establishing practices to control and remove contaminants from drinking water and to prevent its deterioration during storage and distribution; and assessing the nature and controllability of releases of toxic substances to the air, water, and land from manufacturing processes and subsequent product uses. This publication is one of the products of that research and provides a vital communication link between the researcher and the user community.

This report represents the results of an indepth study on the economics of water supply utilizing an econometric technique called the translog cost function. The data analyzed was from a large cross-sectional and historical study of the costs of water supply utility management in the United States. A unique approach was taken by assuming that a water supply utility is a multi-product firm producing both residential and commercial and industrial water. Such issues as the existence of scale economics, marginal cost and second best pricing and the presence of cross-subsidization among customer classes was examined.

Francis T. Mayo, Director
Water Engineering Research Laboratory

CHAPTER 1

INTRODUCTION

1.1 NATURE OF THE PROBLEM

The present water supply situation can be described as substantially different from that of the past. Scarcity of water is evident in many locations in the United States as well as in other countries. Per capita consumption of water has continued to rise with increasing affluence and urbanization, while many existing water sources have become limited in availability and degraded in quality. Increasing energy costs coupled with inflation and high interest rates have raised costs of water provision. In addition, more stringent environmental regulations (e.g., the Safe Drinking Water Act of 1974) have the potential for putting upward pressure on water rates.

In awareness of growing water supply problems, the United Nations has designated the 1980's as the "international Drinking Water Supply and Sanitation Decade". Its primary goal is to provide safe water and sanitary disposal of human wastes. However, in this period of rising costs and impending regulatory changes, the costs of water supply and the factors affecting these costs have taken on new importance.

The conventional solution to water supply problems has been to augment supplies by expanding the system and acquiring the resources as needed to meet the full demand of customers. But the problems of

supplying water of acceptable quality are no longer simple. Issues related to substitutabilities among water supply inputs have become important for the policy maker when he is interested in understanding implications of increasingly scarce resources. Knowledge of scale economies and its limitations have implications for "regionalization" policies.¹ Also, utility commissions have begun to devote more attention to rate structure reform. Public versus private provision of water supply is becoming an important issue as a way to improve management and operation of water supply.

With increasing emphasis being placed on water supply issues, there is a greater need for an in-depth analysis of the cost and economic structure of water supply. Previous research efforts to meet this need have been limited in scope and generality.² The present research seeks to fill that gap.

1.2 PURPOSE OF THE RESEARCH

Recently, considerable research has been conducted in the neo-classical theory of the firm. The main concept employed in this research is the duality principle relating production and cost, as originated by Shephard (1953), together with the flexible econometric specification of production technologies.

¹ Regionalization in the water supply field refers to consolidation or mergers of water utilities. For more on regionalization, see the next chapter.

² A literature review on water supply costs is provided in Section 2.3.

Most theoretical and empirical work in this area has concentrated almost exclusively on single-product firms employing labor, capital, and other inputs such as natural resources, energy, and structures; see, for example, Diewert (1974a, 1978b), Berndt and Wood (1975), Christensen and Greene (1976), Woodland (1976), Fuss and McFadden (1978), Pindyck (1980), Cowing and Stevenson (1981), and Berndt and Field (1981). But, the more recent literature has placed greater emphasis on the multiproduct nature of the firm; see Shephard (1970), Hall (1973), Lau (1972), Diewert (1971), and McFadden (1978). However, empirical studies on multiproduct firms and industries are rather limited.

A few studies analyze the aggregate economy producing two distinct outputs--consumption and investment goods--by employing two primary inputs--labor and capital--together with imported materials; see Christensen, Jorgenson, and Lau (1973), Burgess (1974, 1976), and Denny and Pinto (1978). Other studies suggest the applicability of the methodology of the multiproduct firm to regulated firms and utilities; see Hasenkamp (1976), Brown, Caves, and Christensen (1980), and Denny, Fuss, Everson, and Waverman (1981). Railroads, electric utilities, and telecommunications firms are the most familiar examples of regulated firms.

The aim of this research is to develop a comprehensive economic model suitable for applying the recently advanced neoclassical theory of the multiproduct firm to the analysis of the production structure of water supply. In this study, the water supply utility is viewed as a regulated multiproduct firm offering multiple services to various classes of consumers with spatial and temporal variations. Particular

attention is directed to examination of production characteristics of the multiproduct firm. It is assumed that many issues in water supply can be resolved by applying some basic neoclassical multiproduct production theory to the specification of the technology of water supply by taking into account distinct features of the water supply utility. Nevertheless, the specification and estimation of the technology of water supply employed in the extant research is derived entirely differently from the neoclassical microeconomic theory.³

In the development of the model, the following issues are addressed: (1) restrictiveness of existing specifications of water supply technology in view of the neoclassical theory of production; (2) investigation of input substitution and utilization in the production process; (3) existence of scale economies; (4) examination of marginal costs and the transformation function of water supply; and (5) applicability of second best pricing and presence of cross-subsidization among customer classes due to inappropriate pricing structures.

Although this research is primarily concerned with the analysis of water supply problems, it encompasses many practical problems and issues faced by other utilities and regulated industries. Therefore, the analytical and empirical methodology adopted in this research should provide a useful framework to the study not only of the water supply industry but also of the multiproduct firm in general. Econometric analysis of the general multiproduct firm is far from complete in view of

³ A complete discussion of specification problems in the existing water supply literature is provided in Section 2.4 in the next chapter.

the fact that most of the work in this area has been only preliminary and suggestive without regard to practicality and policy relevance.

1.3 METHODOLOGY OF THE RESEARCH

The approach of this research is to estimate the cost structure of the water supply industry using the recently refined translog function. The translog function can deal with multiple inputs and outputs, variable elasticities of substitution among inputs, variable elasticities of transformation among outputs, and variable economies of scale. The theory of duality is explicitly explored in the estimation of the translog cost function.

The water supply firm is viewed as producing multiple outputs for different customer classes such as residential and nonresidential users. Capital, labor, and energy are considered as the major inputs into the production process for the delivery of water to meet customer demands. Interaction among outputs and inputs are allowed in the estimation, and various hypotheses of the production structure are tested. The data employed represents a cross section of water utilities in the United States for the year 1973. This data were collected for the U.S. Environmental Protection Agency (EPA) primarily by Act Systems, Inc.

An unusual feature of the methodology is the direct incorporation of the firm's operating characteristics into the specification of the technology. Operating characteristics include capacity utilization and service distance which are considered important for production and delivery of water supply. Using the generalized cost function which includes operating characteristics in addition to conventional arguments

of outputs and input prices, it is possible to obtain insights into the quantitative relationship between costs and operating conditions.

In the estimation of the translog cost function, estimates by ordinary least squares would be imprecise because of possible multicollinearity resulting from a large number of regressors interacting one another. Therefore, the iterative Zellner efficient method is utilized which jointly estimates the translog cost function and the input demand functions derived as the first order conditions of cost minimization. The iterative Zellner efficient method is asymptotically equivalent to the maximum likelihood method.

1.4 PLAN OF THE RESEARCH

In this chapter, a brief overview of the issues and methodology has been provided. Chapter 2 presents a general discussion of the water supply industry. This chapter will serve as a background for the model to be developed in Chapter 3. The chapter first examines the characteristics of water supply. Then it reviews the current status of literature on water supply costs. The problems arising in existing specifications of water supply technology are also mentioned.

Chapter 3 is concerned with the development of the model to be empirically tested. The discussion in this chapter is based on the recently advanced theory of duality that exists between transformation and cost functions. The dual multiproduct cost function is derived under the assumption of cost minimization by water utilities. This chapter also discusses criteria of testing the underlying structure of water supply within the framework of the translog multiproduct cost function.

Then comparative static analyses with regard to production characteristics are discussed.

The theoretical model developed in Chapter 3 does not provide any guide to the empirical implementation. Chapter 4 presents the empirical framework under which the model is formulated statistically to facilitate econometric estimation. The chapter explains the estimation method to be utilized and the data chosen for this research. A brief analysis of the sample data is also given in this chapter.

Chapter 5 presents the results from the estimation of the model. The estimated translog multiproduct cost function is examined for flexibility in estimating the cost function for water supply and current specifications that lack a theoretical reference to the neoclassical cost function are critiqued. The chapter also presents various results concerning production characteristics of water supply such as: elasticities of substitution among inputs or outputs, input requirements, scale economies, and marginal costs. In addition, pricing decisions and cross-subsidization among user classes are discussed.

Finally, in Chapter 6, the summary and conclusions based on the findings of the empirical analyses are presented. Policy recommendations and suggestions for future research are also discussed.

CHAPTER 2

WATER SUPPLY INDUSTRY

This chapter provides a general discussion of the water supply industry.¹ Section 2.1 examines the characteristics of water utilities. Section 2.2 discusses current problems the water supply industry is facing. Section 2.3 provides a brief literature review on water supply costs. The literature review is not meant to be exhaustive, but is intended to shed some light on issues later discussed in this research. Section 2.4 considers certain problems arising in past specifications of water supply cost functions as reviewed in Section 2.3. The discussion in Section 2.4 will serve as a motivation for the model which is to be developed in the next chapter.

¹ This chapter, especially Sections 2.1 and 2.2, may be skipped without loss of continuity by readers who are primarily interested in application of the theory of the neoclassical multiproduct firm or who have some knowledge about water supply.

2.1 CHARACTERISTICS OF WATER SUPPLY SYSTEMS

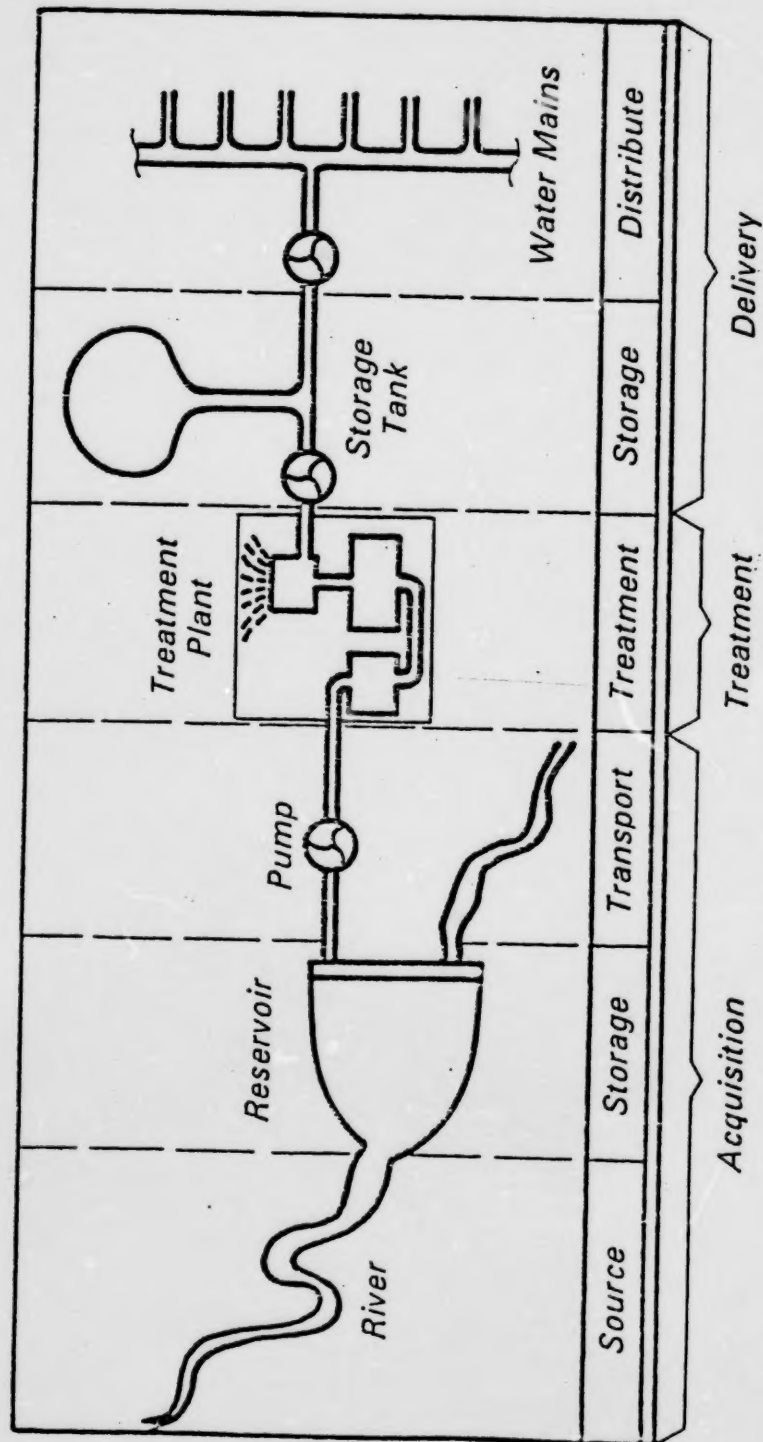
Water supplies provide services or activities associated with the production or acquisition and distribution of water to the public. Water is used by the public for drinking and culinary uses, bathing and washing, heating and air-conditioning systems, watering of lawns and gardens, recreational use in swimming and wading pools, numerous and varied industrial processes, and protection of life and property against fire. Water supply service, like other public utilities (such as electricity, cable television, sanitary services, and telephone services) is often without competition. The local water utility is ordinarily a monopoly except to the extent that industrial customers of water may furnish their own supplies.

Water Supply Systems

Water supply systems are complex networks designed to collect, store, and distribute water to consumers. Water supply systems are physically composed of production facilities and delivery systems. Production systems consist of acquisition and treatment facilities. Delivery is divided into transmission and distribution. A simplified scheme of water supply systems is presented in Figure 1.

Acquisition involves either tapping a source of water that is adequate in quantity to satisfy present and reasonable future demands on a continuous basis, or to convert an intermittent source into a continuous supply by storing surplus water for use during period of low flows. The cost of acquisition includes the collection of raw water from a surface

Figure 1
A Typical Water Supply System



water source (lake or stream) or ground water source (well or spring). Pumping and plant facility cost, labor, and energy are the major portions of acquisition cost.

If water is not of satisfactory quality at the point of acquisition, treatment plants purify it; that is, polluted water is disinfected, esthetically unattractive or unpalatable water is treated to make it attractive or palatable, water containing iron or manganese is subjected to deferrization or demanganization, corrosive water is stabilized chemically, and excessively hard water is softened. The costs of treatment consist of chemicals, sand, power, and depreciation of the physical plant.

Water is transported and delivered to consumers through transmission pipelines and distribution mains. Transmission pipelines are the major truck lines used to transport large volumes of water. They connect the treatment plant to the distribution system. Pumping stations, pipelines, labor, and energy comprise the major costs of transmission. The distribution works include the meters, pipelines, and storage facilities (water tanks) necessary to transport water from the transmission system to consumers. A pumping station might also be involved if ground storage exists. Therefore, distribution systems include the cost of pipelines, water towers, pumping stations, and energy.

Constructing and operating an urban water works system involves large investments in both production (acquisition and treatment) facilities and delivery systems. Each of the two components of a water supply system is usually assumed to have a different cost function. The unit costs associated with acquisition and treatment facilities are generally assumed to

decrease as the quantity of water provided increases. However, the delivery system is more directly affected by the characteristics of the area being served.

Therefore, a characteristic of many water supply and management problems is the tradeoff between the cost of building and operating treatment facilities to meet demand for water and the cost of transportation of water to consumers. If transportation costs outweigh facility costs and the assumed economies of scale associated with water treatment, then decentralization of production to many locations would be economical. Should the opposite be true, centralization of production is preferable; see Beckstein (1975).

Systems Profile

There are several hundred thousand public water supplies in the United States. Community water supply systems serve most of the population. One source of data on water systems in the United States is an Environmental Protection Agency survey (1977) of public water systems. Tables 2.1 and 2.2 present a portion of these data. Approximately 40,000 community water systems were in operation in 1976, serving a total population of about 190 million. The vast majority of the water utilities are small in terms of the number of people served. About 80 percent of the nation's water supplies serve about 8 percent of the total population. However, it is the large systems which are responsible for providing most of the drinking water in the United States. About 8 percent of the water utilities serve about 80 percent of the population. Table 2.1 gives a tabulation of the system size in terms of the number of systems and people served.

Table 2.1
Central Water Systems and Population Served in the United States

Total Population Served	Number of Systems	Percent of Systems	Population Served (millions)
25 - 99	6,308	18	0.4
100 - 499	11,714	34	2.8
500 - 999	4,932	14	3.4
1,000 - 2,499	4,850	14	7.6
2,500 - 4,999	2,496	7	8.7
5,000 - 9,999	1,646	5	10.4
10,000 - 99,999	2,442	7	73.8
100,000 - 999,999	232	1	58.9
>1.0 million	<u>11</u>	<u>—</u>	<u>26.2</u>
Total	34,631	100	192.2

Source: U.S. Environmental Protection Agency. Survey of Operating and Financial Characteristics of Community Water Systems, Washington, DC: Report prepared by Temple, Barker, and Sloane (Wellesley Hills, Massachusetts), April 1977, p. II-4.

In addition to the large number of small community water supplies in the United States, there are quite a number of noncommunity water systems. Most of these systems are privately owned and are found at service stations, motels, restaurants, rest areas, camp grounds, state parks, beaches, national parks, national forests, reservoirs, and other locations frequented by the traveling public. Some schools and industries are also included in this category. Data on these systems are limited.

Most community water supplies are either publicly (government) owned or investor (privately)-owned. Investor-owned water utilities generally operate on a profit basis, subject to state regulations of price and conditions of service. The publicly-owned (mostly municipal) water firm is confronted by control ranging from city council review to independent commission regulation. Although theoretically a non-profit firm, the public system may generate surplus revenues that subsidize local government operation. Public systems dominate the water industry in terms of both absolute number (approximately 85 percent) and service provided, i.e., the number of people served (approximately 99 percent). Publicly owned systems tend to operate in rural areas. The public systems tend to provide more extensive water treatment than the private systems which often provide no treatment whatsoever.

There have been many opinions expressed as to why the water utility industry is dominated by public ownership. It appears that political and social factors have often been given more weight than economic arguments

in the choice of ownership form. There may be three factors leading to public ownership in the water utility industry.

One, water utilities are normally capital intensive. The low rate of return earned on capital in the water utility industry compared with alternative investments discourages private market provision of water and encourages municipalities to develop their own system. Two, public provision of water services is due to the increased emphasis on public health. People become concerned with the quality of water supplies. Today, rigid standards on water quality are imposed at all levels of government. But there appears to be no inherent reason why publicly-owned systems provide any better sanitary quality of service than private firms. Three, there is the importance of fire protection to a city and the difficulty associated with charges for this service.

Water systems obtain their supply from surface, ground, and combination of surface and ground sources. They may purchase water from other water systems. Surface water sources include rivers, lakes, and reservoirs which are generally unsuitable for potable water supply without extensive treatment. Ground water is generally available from wells and from springs in some areas. The quality of ground water varies from water needing only disinfection to water needing extensive treatment for removal of total dissolved solids. An EPA study (1977) shows the predominance of ground water for water systems. About 75 percent of all water systems use ground water. About 12 percent of water systems purchase water and 13 percent of the systems use surface water as a source of water.

The breakdown by system size of some key operating statistics are presented in Table 2.2. One of the items in the Table is the average daily production of systems. The mean production level ranges from .006 MDG (million gallon per day) for systems serving 25 to 100 people, to 496.660 MGD for systems serving over 1 million people.

To operate efficiently, a water utility must be able to generate adequate funds at the proper time. These funds may come from taxes, user charges, or municipal bonds; see Helms and Clark (1978). Tax levies may include property, income, earnings, or special-use taxes. Bonds, either general obligation or revenue, may provide a lump sum to pay for a large project in service over a number of years. Most fixed facilities are financed from bonds of various types or from surplus revenues available from the utility's previous revenue-raising activities. Financial planning for most water utilities therefore entails decisions as to the use of surplus or debt and, if debt is chosen, its maturity structure.

Table 2.3 gives major expenditure categories for government utilities. Capital accounts for 32 percent of all water supply expenditures in Fiscal 1977. Salaries and wages account for 22 percent (39 percent of current operating expenditures). Interest on debt, materials and supplies, energy, purchased water and the like account for the remaining 46 percent.

2.2 PROBLEMS OF WATER SUPPLY AND WATER SUPPLY SYSTEMS

In the past, the supply of water has been considered a free good available in almost unlimited quantities, and the delivery of water to consumers has been considered a routine matter. Thus, the importance of

Table 2.2
Operating Characteristics of Water Systems

Total Population Served	Number of Systems	Average Daily Production (MGD)*	Number of Systems Primary Water Source		
			Ground	Surface	Purchased
25 - 99	6,308	.006	5,930	63	315
100 - 499	11,714	.025	9,488	586	1,640
500 - 999	4,932	.075	3,452	888	592
1,000 - 2,499	4,850	.200	3,347	824	679
2,500 - 4,999	2,496	.480	1,448	674	374
5,000 - 9,999	1,646	.921	1,053	428	165
10,000 - 99,999	2,442	5.059	1,270	757	415
100,000 - 999,999	232	48.003	65	137	30
> 1.0 million	11	496.660	1	10	0
Total	34,631		26,054	4,367	4,210

Source: U.S. Environmental Protection Agency: Survey of Operating and Financial Characteristics of Community Water Systems, Washington, DC: Report prepared by Temple, Barker, and Sloane (Wellesley Hills, Massachusetts), April 1977, p. II-7.

*MGD = millions of gallons daily.

Table 2.3

Finances of Government-Operated Water Utilites
 By Type of Government - 1976/1977
 (Millions of Dollars)

Type of Government	Revenue	EXPENDITURES					
		Total	Capital	Interest On Debt	Current Operations		
					Total	Salaries & Wages	Other
Total	4,995	6,381	2,047	786	3,547	1,395	2,152
States	6	21	3	6	13	8	5
Municipal- ities	3,823	4,306	1,261	475	2,570	1,085	1,485
Special Districts	772	1,413	518	206	688	194	494
Townships	139	187	44	20	124	47	77
Counties	256	453	221	79	153	63	90

Source: U.S. Bureau of Census, 1977 Census of Governments--Compendium of Government Finances, Washington: Government Printing Office, 1979, p. 33.

water supply has been neglected in the minds of the general public and little attention has been devoted to water supply problems. Many factors have contributed to this phenomenon (Mann 1981). Water has generally constituted a relatively small portion of consumer budget and firm production costs. Public health issues have in the past dominated economic and financial concerns. As a result, water historically has been supplied at a lower cost than other utility services. And, the cost and pricing of water services has been ignored as an important variable partly due to the engineering emphasis in water resource development; see Hirschleifer, De Haven, and Milliman (1969). In addition, the water supply industry is characterized by the relative smallness within the United States economy with relatively static technology.

Water may now be no longer free. A shortage and scarcity of water is evident in many locations in the United States and in other countries as well. The problems of supplying water of acceptable quality to an increasingly urban and decentralized population are no longer simple matters. Spreading urban boundaries force many potential water consumers to locate farther and farther away from available sources. Some areas that are inherently water-limited have attracted significant population growth, thereby straining the available water resources. A scarcity of land, labor and capital needed to transport water to its point of use has contributed to these problems.

Furthermore, the recent high rate of inflation and increasing costs have produced increased labor, treatment, pumping, and construction costs, thus resulting in more burden to utility managers and in higher frequency of water rate increases than in the past. Rising energy

costs may dictate significant capital investment in the future to increase energy efficiency. In addition, the increased environmental regulation in the water supply industry resulted in increased treatment and purification costs.

Recent passage of the Safe Drinking Water Act (SDWA) of 1974 (Public Law 95-523)² has intensified a growing concern and interest in problems related to water supply and water supply utility management and in examining the ways in which drinking water is handled before it is delivered to consumers. The purpose of the Act is to assure that the public is provided with safe drinking water and that water supply systems serving the public meet minimum national standards for the protection of human health. The Act will bring about fundamental changes in the way in which regulatory agencies, utilities, and the public view the problems of delivering water to consumers.

The Act authorizes the U.S. Environmental Protection Agency (EPA) and state agencies to develop and enforce water quality standards, and it establishes a mechanism for assuring compliance with the standards. The Act may have a significant impact on the cost of supplying water and it may further compound the problems of small water systems. Water systems in general may be forced to install new technology in order to meet these standards. But small systems will have a potentially serious problem. High unit costs are generally associated with small systems and if small systems are forced to install expensive treatment technology, it is conceivable that per capita costs could more than double.

² See Clark (1979) and Clark and Stevie (1978) for a good discussion of the Safe Drinking Water Act.

Thus, problems related to water supply and water supply systems are of increasing concern in the United States. Small water systems in general have more problems than large systems. One major problem for small systems is their lack of operating revenue. They cannot benefit from economies of scale as do large urban systems because they are small in terms of the number of connections served. Certain types of services must be provided, no matter how few the number of connections, causing the unit costs to rise as the number of connections being served decreases. The initial construction cost of the treatment and distribution system may also be high relative to the number of connections.

Another major problem associated with small water supplies is that of providing high quality drinking water. Small systems tend to have more quality control problems than do large systems. Epidemiological evidence has shown that the majority of recent waterborne disease outbreaks occurred in small systems (Clark 1979). These outbreaks are primarily due to the inadequacy of treatment technology, which is of course related to inadequate revenues.

A commonly suggested panacea for small systems problems is "regionalization", in which a small community can join with other small communities or with a large community to form a regional water system. Two approaches can be considered when attempting to develop a regional system. One approach involves the centralization of purchasing, maintenance, and various management support items. Another approach involves physically interconnecting existing systems.

Regionalization provides an efficient means of internalizing many of the costs associated with water supply. The basic motivations for

regionalizing water systems are: large systems motivated by exhaustion of sources within their political boundaries and small systems motivated by high costs of adjusting to new drinking water standards. The alleged justifications or benefits with regionalization are:³ (1) the existence of economies of scale with regional systems which may result in decreased unit costs to consumers; (2) increase in quality of services, both reliability and quality of supplied water; and (3) better water management within the region, regarding allocation of water to several different types of users and environmental protection.

Thus, increasing quality requirements and associated increases in costs will no doubt increase interest in the regional concept particularly among small systems. Consequently, the Safe Drinking Water Act, along with other environmental legislation, encourage the development of regional systems for managing water supplies to offset the increased costs through scale economies and other gains in economic efficiency.

Regionalization has great appeal to regulatory groups because it reduces the number of individual entities with which they must work. However, traditional engineering interpretation of regionalization has often resulted in the construction of large central treatment plants, thereby minimizing treatment costs with little or no consideration given to the costs associated with the transportation and distribution of water. Further, each argument favoring regionalization of water supply systems is not conclusive by itself, since special conditions for each

³ See the American Water Works Association (AWWA) Committee on regionalization (1979, 1981).

case may cause the advantage not to exist, as discussed by Clark (1979b, 1980).

2.3 RELEVANT LITERATURE REVIEW⁴

Numerous studies have focused on the determinants of water costs and cost behavior, but results have not been entirely satisfactory. These studies have usually consisted of regressions of costs against output, with a number of variations in functional form and differences in independent variables included. Common to these studies is, however, the functional breakdown of water supply costs into five system components: acquisition, treatment, transmission, distribution, and overhead.

Previous research has been primarily concerned with investigating scale economies. This is due mainly to the fact that the direction of public policy toward the water supply industry heavily depends on the nature and existence of scale economies in this industry. Within the water supply industry, evidence of scale economies is central to the continuing debate over the existence of natural monopoly, the applicability of marginal cost pricing, and the possible advantages of increased competition.

A typical functional form widely employed in these studies is the simple regression relating costs to design capability:

$$C = aY^b$$

⁴ For a complete review of literature on water supply costs, see EPA report (1978) and Mann (1981).

where C is the system costs, Y is design capacity as measured by the actual water produced, a and b are constants to be estimated. There are economies of scale if b is less than 1; diseconomies of scale if b is greater than 1; and constant returns to scale if b is equal to 1.

In one of the earlier studies, Orlob and Lindorf (1958) in the estimation of water supply cost of treatment found that the unit cost decreased as the quantity of water treated increased. Their studies were extended by Konig (1967) and Hinomoto (1971) by incorporating a capacity utilization variable. In a study of water transmission costs, Linaweaver and Clark (1964) found monotonically decreasing unit costs as larger quantities of water were piped through larger pipe sizes. Hines (1969) examined long run cost functions for relatively small water systems concluding that water systems using surface sources experience decreasing unit costs with increasing size. Ford and Warford (1969) found that increasing per capita consumption and/or increasing consumer density within a service area tended to decrease unit costs. And, Andrews (1971) found that substantial economies of scale in water service occurred with increasing per capita usage but not necessarily with increasing service connections.

Hanke and Davis (1973) cited evidence of increasing unit costs in water provision and the virtual exhaustion of economies of scale for some large systems. Also, a number of EPA studies⁵ indicated that while economies of scale might exist in the treatment process, the unit cost of

⁵ See Clark (1975, 1979), Clark, Gillean, and Adams (1977), Clark, Stevie, and Trygg (1976), and Clark and Stevie (1981).

water supply increased with increasing distance from the treatment plant and decreased with increasing population density within a specific area. Therefore, economies of scale in water treatment may be offset by diseconomies of scale in water transportation and distribution.

In summary, most existing studies agree on the existence of economies of scale in water treatment; some engineering estimates indicate unlimited economies of scale in water treatment. Estimates of distribution and other components of water supply cost functions are not in general agreement. However, empirical estimates of the overall scale economies for each utility's complete production activity often suffer from the lack of overall comparability of component system scale economies. The inherent differences of system components makes it impossible to make direct comparisons of component scale economies.

It is generally believed that the mode of ownership has different impact on the cost behavior of the firm. In this regard, the common notion is that public ownership is inherently less efficient than private ownership. Studies on the economic efficiency of public and private ownership in water utilities are rather limited. Mann and Mikesell (1976) using cross sectional data found that private water firms had significantly higher operating costs than the public water firms. They attributed part of this difference to the higher compensation paid to private employees. Another reason for the difference was the effect of regulation on operating performance. While admitting that their study provided no definite answers on comparative efficiency of private and public water utilities, Mann and Mikesell concluded that there was no economic justification for consolidating water utilities to

achieve economies of scale, since their output variable indicated diseconomies of scale in larger utilities.

The metering and pricing of water service is a widespread phenomenon in the United States, if not in other countries. Pricing can have significant impact on the demand for water. It can also have substantial effect on the requisite capital investment in water supply facilities. Thus, it may be appropriate to ask which is the appropriate rate structure, especially with respect to delivering water to consumers under varying circumstances. Marginal cost pricing is strongly advocated as the appropriate rate of water services.

The common practice in water supplies is based on the average historical or embedded costs instead of marginal replacement costs. Further, it employs one rate schedule for all customers with the single schedule incorporating consumption blocks reflecting the use and cost characteristics of each user class; see Keller (1977). These current pricing practices may result in misallocation of resources to water supply development and cross-subsidization.

Misallocation of resources occurs in the form of over-or-under-investment of capital (EPA report 1978). Cross subsidization occurs because the marginal cost of supplying water varies over many dimensions, among which are quantity treated, customer class, distance, and time. Then, the existing uniform water pricing can lead to cross-subsidizations of users among customer classes (residential and non-residential), between those near and those distant from the treatment facilities, and between peak and off-peak periods (Hanke 1972). Reiter (1977) emphasized that interfirm comparison of water prices were

difficult due to variances in source of supply, topography, climate, system age, customer density, consumer mix, water quality, service area growth, inflation rates, energy costs, service volume, per capita usage, and financing of fire protection facilities.

A number of water rate reforms have been proposed in the literature. Howe (1968) suggested a major reform in the rate design process in that it does not recognize the interrelationship between demand, system design, costs, and rates. Hanke (1978) advocated similar reforms in proposing an integrated engineering-economic approach to rate setting and capacity planning. Both Hanke (1972) as well as Hanke and Davis (1973) advocated seasonal peak load pricing (based on marginal cost) arguing that uniform rates over time results in allocative inefficiency, an involuntary subsidy to peak users by off-peak users, and the inducement of system expansion beyond efficient sizes. Feldman (1975) advocated water rates reflecting peak hour and peak day water demand to complement seasonal pricing. Hanke (1975) as well as Hanke and Davis (1973) also argued against uniform water pricing over space. That is, conventional uniform water rates throughout the service area result in the subsidization of low-density customers by high-density customers, overcapitalization of water systems, and urban sprawl.

2.4 PROBLEMS OF EXISTING STUDIES⁶

In the preceding section, the literature on water supply was briefly reviewed which is relevant to the issues to be discussed in this study. While numerous studies have been performed in the estimation of water supply costs, these studies are limited in scope and generality; unsophisticated in their use of methodology. Most importantly, the existing specification of the water supply cost function is entirely different from the neoclassical cost function. A proper specification of the cost function as we know it from the neoclassical theory of the firm requires inclusion of output variable and input prices as its argument. But determinants of costs for water supply employed in the previous empirical studies are arbitrary and crudely ad hoc with no theoretical reference to the underlying technology of production.

A number of empirical studies concentrated on the relationship between costs and output, neglecting the role that input prices play in cost function. In fact, no one has yet hit upon the idea of including input prices explicitly in the statistical cost function of water supply. An implicit assumption behind this specification of excluding input prices may be that input prices are constant and independent of the purchases by a firm. But actual observations frequently come from successive time periods during which input prices may have changed substantially in response to influences other than the firm's purchases.

⁶ For a full understanding of this section, it is advisable to first read the discussion of the specification of the technology of water supply in Section 3.1.

In fact, in recent years, the relative prices of capital, labor, energy, and materials in the United States have been altered radically as a consequence of the increase in the price of energy relative to the prices of other productive inputs. Substitution possibilities among various inputs of production require the knowledge of input prices. But due to the failure of inclusion of input prices in the cost function, previous studies were not able to address the issue of substitutabilities among water supply inputs.

Furthermore, most previous studies fail to recognize the fact that the output of a water supply firm is multidimensional by its nature. The water supply firm provides water to various classes of customers at various times as well as over various distances. Then production in the water supply industry is best characterized as a multiple output production process. Recognition of this dimension is important to reflect differences in the cost of providing water for various uses. Therefore, it is clearly inappropriate to estimate the cost function by using a single aggregate measure of output. While a few earlier studies recognized the multiproduct production process of the water supply firm, these studies employ specification of technology different from neo-classical theory which can employ restrictive assumptions such as distinct production process for individual outputs (nonjoint production) or separability of outputs from inputs. For example, an EPA study (1978) specifies a separate cost function for residential and nonresidential supply of water. Unfortunately, this implies that there is no overlap of facilities in production, a very restrictive specification.

In addition to the inappropriate specification of the structure of

technology, existing studies employ the extremely simplified functional forms such as linear or linear logarithmic. These functional forms, along with Cobb-Douglas or CES forms, are restrictive for use in modelling the cost of water supply and the economic structure of production. They impose certain a priori and unjustified restrictions on characteristics of the production process such as scale economies and elasticities of substitution. For instance, because of the linear logarithmic forms employed in the previous studies, scale economies are assumed to be constant for all levels of output, contrary to our expectation. The use of extremely restricted functional forms, together with inappropriate specification of the production technology, may have resulted in substantial errors in the implied estimates of scale economies in the existing studies.

Last but not least, it is important to consider operating characteristics that are important for production and delivery of water supply. Even if two water firms provide similar services to consumers, they may differ widely with respect to operating conditions in treatment and spatial dimensions in the distribution systems. Consequently, in analyzing the cost function, considerable care should be used to incorporate the effects of the operating characteristics where appropriate. However, the specification of operating and spatial variables in the existing studies has been arbitrary and no clear knowledge about these variables has been obtained.

Knowledge of the nature of the cost functions for water supply is essential for both policy analysis and the utility's business decisions. Previous studies fail to estimate the cost function that appropriately

characterize the structure of water supply, and, as a consequence, the validity of the empirical results remain questionable. In the next chapter, we propose a complete formulation of a costing methodology for water supply based on the neoclassical theory of the multiproduct firm. By working with the model in a manner that is entirely different from those that have been used before, we can better understand the structure of water supply and analyze various issues facing regulators and water supply managers.

CHAPTER 3

FORMULATION OF THE MODEL

In the preceding chapter, a general discussion of the water supply industry was presented. Now, a detailed exposition of the theoretical model is given which will form the appropriate framework of empirical analysis in water supply in the next chapters. The chapter begins with the development of the multiproduct cost function for the water supply industry under the assumption that water utilities minimize cost for a given level of outputs. The duality principle is explicitly utilized in the development of the multiproduct cost function. In Section 3.2, the translog form is used to specify the functional form for the cost function. Properties of the translog cost function are also discussed.

Section 3.3 presents tests of various restrictions on producing water supply implied by the translog cost function. These tests include homogeneity of outputs, input-output separability, and nonjointness in production. Sections 3.4 through 3.8 discuss various comparative static (sensitivity) analyses with regard to production characteristics of the water supply firm under the framework of the translog cost function. These sections constitute the main ingredients of this research. The input side of water supply is discussed, followed by the cost or supply side. Pricing decisions in water supply are also discussed.

3.1 CHARACTERIZATION OF THE COST FUNCTION FOR WATER SUPPLY

The U.S. water supply industry is a regulated public utility¹ providing services associated with the production and distribution of water to the public. Municipal water utilities, which are the main concern in this study, account for nearly 85 percent of the total supply of delivered water. The technological and institutional characteristics of the water supply industry which are important for the model to be developed are as follows:

1. Water supply systems require massive amounts of capital in the form of plant and related equipment prior to commencement of services. The systems also require extensive investments in transmission pipelines, pumping stations, and storage facilities in the distribution works which are necessary to convey the water to consumers. Capital costs are in general related to the design capacity of the plant and not affected by the actual level of production of water. Therefore, in a given time period, marginal costs of capital may be zero, but as the system expands, costs would increase.

The capital market in which utilities seek funds for expansion is highly competitive and the rates at which individual utilities can borrow funds are little affected by individual actions over a wide range. Construction costs vary geographically and also appear to be unaffected by an individual utility's actions.

¹ Stevie and Clark (1982) note two types of regulations imposed by various agencies on water utilities: regulations on profits and quality of water. In general, the extent of regulations varies whether utilities are investor-owned or publicly-owned (municipal).

2. Wages and salaries comprise a significant portion of the utility's budget. These include payments to operators, foremen, managers, and office clerks, for example. Other payments such as insurance and retirement plans may be included in labor expenditures. Most labor payments vary with the actual level of production of water. However, wage rates are set by contracts that extend over a period of time. Over long periods, wages appear to be determined competitively.

3. Treatment of water requires heavy use of energy. Energy is also required in the transportation of water over rough terrain and long distances. The energy cost of pumping water depends on the quantity of water supplied, the distance to a demand point, and the net altitude between the demand point and supply point. The price of energy is determined in competition with other uses. Other costs of supplying water to consumers include chemicals (used to transform the raw water into a quality suitable for consumption) and maintenance materials. But these costs are a relatively small portion of the utility's budget.

4. The water firm provides various services to markets having dissimilar demand characteristics. Demand characteristics refer to various dimensions of water demanded by users over time, across customer class, and over space in various quantities. The demand for water, as a function of time, refers to peak demand. The demand for water by customer class distinguishes residential, industrial, commercial, and other users. Spatial aspects of demand refer to the form and structure of the areas and neighborhoods being served. Spatial variables include service distance, population density, and topography, but service distance appears to be most directly related to the distribution

activities of water supply.²

Marginal costs also change for the various uses of delivered water. Peak load costs differ from off-peak costs because of the added costs required for the installation of storage tanks, greater capacity pumping stations, and larger sized pipelines. Marginal costs of water vary across customer class. The cost of supplying water to residential customers appear to be higher than to nonresidential users because of higher pipeline, energy, and billing costs associated with residential supply. The cost of providing water to spatially distant customers is higher than to customers located near treatment facilities because of the additional costs to be incurred as the service distance increases.

It appears that, although the water supplied to various users might be physically similar, it can be considered an economically unique output with distinct demand and supply conditions. Nevertheless, different outputs of water are produced in large part under conditions involving joint or common costs typically incurred on behalf of all which cannot be properly assigned to any output separately. The firm's total cost includes some outlays attributable directly to particular outputs in the firm's product line and some expenditures which are incurred in

² This conclusion is based on the following empirical results. An EPA study (1978) found the topographical variable as measured by altitude insignificant. Mann and Mikesell (1979) found the population density variable insignificant. Mann and Mikesell's result was confirmed in our preliminary analysis. The inclusion of population density had not only a wrong sign but also implications with regard to input requirements contrary to our expectation. The obvious reason found later is the fact that density and distance are inversely related especially for residential service: the users who live farthest from treatment facilities, in general, also live in low-density neighborhoods; see Hanke (1972).

common for several outputs simultaneously. For example, the same physical plant is used to provide service during periods of peak and off-peak demand. The same facility is used to provide service to different customer classes. The same system is also utilized to supply water to customers located at various distances from the sources of water.

5. Water utilities are required to provide services in sufficient quantities upon demand by the public. This means that water must be supplied more or less instantaneously and continuously. Consequently, the demand for water by the public dictates both the magnitude and schedule of production of water. The requirement to serve the public forces utilities to be concerned about existing capacity levels and projections of future demands. This often leads to building capacity in excess of current needs in anticipation of a growing demand.³ The rationale for this is that if there are economies of scale (as is usually assumed) and a growing demand, then the utility may find it profitable to add more capacity than they expect to use in the immediate future.

However, much of the capacity of water supply systems lies idle due

³ It is of interest to note other reasons for excess capacity. One is the effect of regulation on capacity utilization. Since Averch-Johnson (1962), there has been some acceptance of the idea that regulation designed to secure some rate of return on capital of the firm will induce excess capacity. See for example Petersen (1975), Spann (1974), and Courville (1974) for empirical tests of the Averch-Johnson effect.

Another is the presence of excess capacity in the monopolistically competitive market in the Chamberline model. Also, Willig (1979) and Panzar and Willig (1981) noted that the firm may leave excess capacity in some stage of production to exploit economies of scope (joint production), because inputs that are shared (or utilized jointly without complete congestion) by the processes utilized to produce several different outputs may be imperfectly divisible. Further, Murray et al. (1978) found for electric utilities that in a period of rising real incomes and rising real electricity prices, electric utilities often faced deterioration in capacity utilization.

to seasonal and hourly fluctuation in demands, although the treatment and pumping facilities must have a high capacity to directly satisfy the maximum demands by the public. Nevertheless, a smaller capacity may be sufficient by utilizing the system more fully and storing excess water accumulated during slack periods in distribution tanks or reservoirs to be used when demand is peaking. Then, capacity utilization reflects both the load characteristics of the utility and the level of reserve margins.

Underutilization of capacity, which is characteristic of a water supply system, is a chronic problem. Where there is a low rate of capacity utilization, there is substantial unused capacity and thus facilities are less efficiently utilized, thereby increasing capacity costs per unit of output. Capacity costs are increased due to additional interest, depreciation (because of wear and tear on capital equipment), maintenance, and labor costs to be incurred because of the increased system size. However, a larger rate of capacity utilization results in more efficient use of facilities (less idle capacity) and is consistent with lower unit capacity costs (less capacity needs). Therefore, the cost structure of water supply will be significantly affected by the extent of capacity utilization.⁴

6. Water utilities are required to supply all the water demanded at regulated prices. Thus, revenues from the sale of water depend primarily on rates set by utility commissions and other regulatory bodies.

⁴ DeVany (1975) argues that it is not necessarily inefficient to have some unused capacity, because, given any reasonable stochastic properties of demand, capacity serves to reduce waiting time and planning costs for consumers.

From the above characteristics of the water supply industry, some conclusions may be drawn which lead to the model developed below. First, it is plausible to view water utilities as producing and selling multiple outputs supplied for various purposes over time, across customer class, and over space. This assumption critically depends upon the fact that demand and production conditions of water vary across outputs.

There are two basic advantages of the multiproduct specification for water supply. One, the difficult problem of forming a consistent aggregation index of the multiple outputs familiar in single product models can be avoided. When the water supply firm provides different services of water under varying circumstances, it may not always be possible to represent aggregate output by using a single variable. The use of the multiproduct formulation eliminates the need for an aggregate measure by disaggregating outputs.

Two, the multiproduct framework provides marginal cost estimates for each output. This information is important in rate structure analysis. Various rate proposals in water supply are based on the notion of allocation of costs.⁵ However, when there are costs of production which are shared by two or more services in the production process, it is impossible to identify cost with individual service in an unambiguous manner.⁶ Since the multiproduct representation of water supply provides marginal costs for each output, we can avoid any need to allocate joint

⁵ See the American Water Works Association, Water Rates Manual.

⁶ It is due primarily to the fact that average cost as defined for the single product case is not defined for the multiple outputs. For various cost concepts for the multiple product case, see Baumol, Panzar, and Willig (1982).

costs and thus can base the rate design on marginal costs as frequently advocated.

While few argue that passenger and freight services are identical outputs in rail or air transportation, most would agree that establishing the difference between services offered, for example, to residential and nonresidential customers of water supply would require empirical investigation. Nevertheless, the characterization of water utilities as providing multiple outputs is not an unusual assumption made in the literature of public utility and welfare economics. For example, in Faulhaber (1975), the water supplied to different neighborhoods is considered a distinct output in the determination of cross-subsidization among users. It is one of the purposes of this research to verify statistically the nature of the multiple output production process of water supply. Due to paucity of data, peak load characteristics will not be considered in this research.

Second, it is possible to view labor, capital, and energy as the major inputs in the production of delivered water. This view relies on the assumption that labor, capital, and energy are as a group weakly separable from other inputs to production such as chemicals and maintenance materials. Weak separability means that the marginal rate of substitution between any two of the first three inputs is independent of changes in the quantity of chemicals and maintenance materials used as inputs (Blackorby, Primont, Russel 1980, Berndt and Christensen 1973, Denny and Fuss 1975, Russel 1975). Lossely stated, this separability implies that the cost-minimizing choice of the mix of labor, capital, and energy, is independent of either levels or mix of chemicals and

maintenance materials. It is important that we assume separability exists in order to reduce our model from all the inputs to only those that we feel are important to the decision making process. Further, this assumption is made necessary by the fact that there are no data available from which to construct price indices of chemicals and maintenance materials.

Third, variations in the rates of capacity utilization will have a significant impact on the measurement of production characteristics of the utility such as input usages and scale economies. Since utilities are extremely capital intensive, relatively small differences in capacity utilization rates can result in substantial differences in required capital inputs. The consumption of energy can also be affected by successive changes in the rates of capacity utilization. Further, since utilization of capacity has a significant effect on the cost of water supply, the inability to account for capacity utilization will bias the measure of scale economies. Therefore, the capacity utilization variable has to be included directly in the specification of the model.

Fourth, despite the fact that the water supply industry does not operate in perfectly competitive markets, it is plausible to regard the outputs of a utility and the prices it pays for productive inputs as exogenous. Capacity utilization is also assumed to be exogenous to the utility. The assumption of exogenous input prices is consistent with the suggestion of Berndt and Wood (1975) that "at the level of an individual firm, it may be reasonable to assume that the supply of inputs is perfectly elastic, and therefore input prices can be taken as fixed."

Formally, the water utility's multiproduct production process can be represented by the transformation function of the form:

$$T(Y_R, Y_N; X_L, X_K, X_E; Z_U, Z_D) = 0^7 \quad (3.1)$$

where Y_R and Y_N are the amount of water delivered to residential and nonresidential users; X_L , X_K , and X_E are labor, capital and energy used in production and distribution of water supply; Z_U is the rate of capacity utilization; and Z_D is distance from the treatment plant to the service area. It is assumed in the transformation function that the utility has available a continuum of technologies, each corresponding to different values of capacity utilization and service distance. The transformation function (3.1) will be compactly expressed as $T(Y, X, Z)$ with Z referred to as a vector of "operating" variables consisting of capacity utilization and service distance.

The transformation function $T(Y, X, Z)$ is assumed to satisfy certain regularity conditions, among which are:

(1) Differentiability: $T(Y, X, Z)$ is continuous and twice differentiable with respect to Y , X , and Z . This condition implies the existence of a continuous and smooth production technique.

(2) Monotonicity: $T(Y, X, Z)$ is monotonically increasing in X and decreasing in Y for a given Z . This indicates the presence of the input-

⁷ Alternatively, it can be presented by:

$$T(Y_R, Y_N; X_L, Z_U X_K, X_E; Z_D) = 0 \quad (3.2)$$

where $Z_U X_K$ reflects the utilized capital.

output relationship such that a further increase in any output (say residential water) requires an increase in some input quantities or a decrease in some other output (say nonresidential water).

(3) Concavity: $T(Y,X,Z)$ is concave in X for given Y and Z . This condition suggests that isoquants are convex to the origin. Taken together, these regularity conditions ensure that the technology of water supply will be nicely behaved.

Many studies in electric and water utilities show that cost minimization is not an inappropriate assumption for municipal utilities; see, for example, Merlove (1963), Neuberg (1977), and Christensen and Greene (1976, 1978). A logical step in this study is to assume that a water utility minimizes the total costs of producing the given outputs by choosing optimal input quantities, subject to the transformation function and the prices it must pay for inputs. Therefore, the utility's task is to:

$$\min C = W_L X_L + W_K X_K + W_E X_E \quad (3.3a)$$

s.t.

$$T(Y_R, Y_N; X_L, X_K, X_E; Z_U, Z_D) = 0 \quad (3.3b)$$

Where W_L , W_K , W_E are the prices of labor, capital, and energy, respectively, and C is the total cost of providing water to residential and nonresidential customers.

Optimization problem (3.3) provides the following system:

$$\frac{W_j}{W_q} - \frac{\partial T(Y, X, Z) / \partial X_j}{\partial T(Y, X, Z) / \partial X_q} = 0 \quad (j \neq q = L, K, E) \quad (3.4a)$$

$$T(Y, X, Z) = 0 \quad (3.4b)$$

of three equations which may be used in estimating $T(Y, X, Z)$. If equations (3.4) are assumed to be associated with a well behaved minimum of (3.3a), a solution for X_j ($j = L, K, E$) exists as functions of outputs Y_i ($i = R, N$), input prices W_j ($j = L, K, E$), and operating variables Z_k ($k = U, D$). Given fairly general functional specifications for $T(Y, X, Z)$, however, it will generally not be possible to express the X_j as explicit functions of Y_i , W_j , and Z_k . To surmount this problem, one can make use of the theory of duality.

The theory of duality between cost and production⁸ ensures that for every transformation function of the type shown in (3.1), there exists a dual multiproduct cost function of the form:

$$C = C(Y_R, Y_N; W_L, W_K, W_E; Z_U, Z_D) \quad (3.5)$$

The cost function (3.5) will be compactly referred to as $C = C(Y, W, Z)$. If the cost function (3.5) is minimized with respect to input prices W_j ($j = L, K, E$), which are predetermined, the principle of duality also ensures

⁸ Duality theorem in production theory was originated by Shephard (1953, 1970), and further developed by Uzawa (1963), McFadden (1978), and Diewert (1974a, 1978b), among others. Nerlove (1963) was the first who applied the theorem in empirical work. Applications of duality theorem to multiproduct firms and industries include Christensen, Jorgenson, and Lau (1973b), Burgess (1974, 1976), Denny and Pinto (1978), Brown, Caves, and Christensen (1980), and Denny, Fuss, Everson, and Waverman (1981).

the existence of the transformation function (3.1) dual to the multiproduct cost function (3.5).

The multiproduct cost function (3.5) has the following properties implied by the earlier assumptions of the transformation function.⁹

(1) Differentiability: $C(Y, W, Z)$ is continuous and twice differentiable with respect to Y , W , and Z .

(2) Monotonicity: $C(Y, W, Z)$ is monotonically increasing in W for given Y and Z . Further, $C(Y, W, Z)$ must be increasing in its outputs Y , particularly if the possibility of free disposal is assumed.

(3) Concavity: $C(Y, W, Z)$ is concave in W for given Y and Z .

(4) Linear homogeneity in input prices: $C(Y, W, Z)$ is positively linear homogeneous in W for given Y and Z . That is, for the fixed levels of outputs and operating variables, total cost increases proportionally when all input prices increase proportionally. This additional condition is tantamount to the absence of "money price illusion" in the minimization of cost, which is a virtually compelling theoretical restriction.

⁹ A given property of the transformation function does not necessarily imply the same property of the dual cost function and vice versa. Rather what is needed are the conditions on the input requirement set:

$$L(Y) = \{ X \mid T(Y, X, Z) \geq 0 \}$$

However, to bypass the complex nature of the duality theorem, the correspondence between the transformation function and the cost function is loosely stated.

Shephard's Lemma (Shephard 1953) assures us that there is a set of input demand equations which can be derived from the cost function.¹⁰ That is, the derivatives of the cost function with respect to a given input price equals the demand for that input:

$$\frac{\partial C(Y, W, Z)}{\partial W_j} = X_j(Y, W, Z) \quad (j = L, K, E) \quad (3.6)$$

Thus, demand for each type of input is a function not only of the price of that input but also the price of other inputs, the outputs, and the operating variables of the firm.

$$X_L = X_L(W_L, W_K, W_E; Y_R, Y_N; Z_U, Z_D) \quad (3.7a)$$

$$X_K = X_K(W_L, W_K, W_E; Y_R, Y_N; Z_U, Z_D) \quad (3.7b)$$

$$X_E = X_E(W_L, W_K, W_E; Y_R, Y_N; Z_U, Z_D) \quad (3.7c)$$

When the transformation function is differentiable in outputs Y , The following condition holds:

$$\frac{\partial C(Y, W, Z) / \partial Y_R}{\partial C(Y, W, Z) / \partial Y_N} = \frac{\partial T(Y, X, Z) / \partial Y_R}{\partial T(Y, X, Z) / \partial Y_N} \quad (3.8)$$

¹⁰ Note that the input demand function is of the sort called a compensated (or Hicksian) demand function, since it is derived for the values of output Y held constant. That is, it is the demand function from which the output effect has been removed so that it describes only the substitution effect. It is distinguished from ordinary (or Marshallian) demand functions which include the demand effect caused by the change in the price of the output. Hanoch (1975a) uses the indirect production function approach to derive the ordinary input demand equations. For applications of the ordinary input demand functions, see Phillips (1978) and Chambers (1982).

That is, the ratio of the marginal costs of two outputs, residential and nonresidential water supply, is equal to the marginal rate of transformation between them. This implies that the transformation frontier of the firm can be derived from the estimated multiproduct cost function.

The significance of the duality theorem is that the transformation function (3.1) and the cost function (3.5) provide the same information about the firm's technology, so the structure of water supply can be examined from either function. The choice of which function to estimate must then be based on statistical convenience and computational advantages as well as analytical purpose. Since the purpose of this research includes the measurements of scale economies and marginal costs for multiple outputs, estimation of the cost function is the most direct way to proceed. Estimation of the cost function is also favored on computational and statistical grounds.

Estimating the transformation function requires specifying the inputs in terms of quantities, while the cost function uses input prices. One of the difficulties with the estimation of the transformation function is the lack of reliable, if available, data on capital stock. The principal problem stems from the fact that the stock of capital is composed of various types and vintages of machines, buildings, and inventories. Combining them into a single measure--or a few measures--is a formidable problem. Consequently, the measurement of capital stock has been an extremely difficult concept to handle properly, either theoretically or empirically.

Another difficulty cast in terms of statistical properties is that firms choose input levels as decision variables. Thus, observed input

levels are endogenous variables and thus correlated with the disturbance in the transformation function. As a consequence, the true relationship between inputs and outputs will be obscured by simultaneity bias. This source of bias can be avoided by estimating the cost function instead of the transformation function. As indicated earlier, input prices and outputs can be treated exogenous for the water supply firm. Therefore, the right hand variables of the cost function are all exogenous and uncorrelated with the disturbance. Further, this research assumes that a plausible model of behavior of the water supply firm is cost minimization. This suggests that estimation of a transformation function for water supply from data of inputs and outputs is incorrect. A correct procedure is to use the cost function. This is the step taken in this research.

Cross section data are utilized to analyze the structure of water supply. It is commonplace to presume that cross section data are likely to yield estimates of the long run cost function, while time series data tend to generate estimates of short run responses.¹¹ The rationale for this assumption rests on the proposition that firm to firm variations tend to be much greater in the cross section samples than in the time series data and that these differentials tend to persist for substantial time periods, so that adjustment to a long run equilibrium is possible.

¹¹ Stapleton (1981) examines cross section, long run, and the time series, short run argument, with regard to elasticities of substitution, and concludes otherwise. According to Stapleton, reliable estimates cannot depend only on whether the data base is cross section or time series but also on the model specification. This issue is beyond the scope of this research, but more will be said in the discussion of elasticities of substitution in Section 5.2.

Therefore, the multiproduct cost function (3.5) is assumed to reflect the long run behavior of the water supply utility.¹² Further, since the observations are taken from various firms in a single year, it is usually assumed that all firms in the sample have access to the same choices in plant design and thus face the same technology. However, it is quite reasonable to suspect that there are differences in the structure of the technology of water supply that cannot be accounted for solely by differences in input prices, outputs, and operating variables that attempt to measure the nature of water supply operations.

In fact, U.S. water systems are characterized by extreme diversity (Mann 1981). Some of the differences in water utilities reflect ownership or organizational arrangements; other sources include differences in regulatory environment, financial markets, management quality, system age, water quality, climate, distribution characteristics (e.g., terrain), source of supply, system scale, size and density of service area, and demand characteristics (e.g., the magnitude, diversity and timing of peak demands). Because of this diversity, it is difficult to obtain the overall comparability of water systems. While dummy variables are often introduced to capture the relevant technological differences among firms, the paucity of the relevant data makes it impossible to undertake an analysis incorporating all the variations of technology of water systems.

¹² Perhaps, a more adequate description of the adjustment behavior of the utility is "quasi" long run, to the extent that the capacity utilization variable included in the model controls for the effect of short run variations in capital and the effect on the long run capital position.

3.2 TRANSLOG SPECIFICATION OF THE MODEL

The conceptual framework for the structure of water supply in terms of the dual cost function developed in the previous section has been entirely general, but the empirical implementation requires specification of a functional form for $C(Y,W,Z)$. It is desirable to specify a functional form which is "flexible," in the sense that no a priori restrictions are placed on the nature of the technology or on the first and second-order derivatives of $C(Y,W,Z)$. It is also desirable to have a functional form which is amenable to interpretation and econometric estimation.

In recent years, several flexible functional forms have been proposed, because of the failure of conventional restricted forms such as Cobb-Douglas and CES (Constant Elasticity of Substitution) to appropriately characterize the structure of production and cost. The best known flexible functional forms are the transcendental logarithmic function (translog for short) proposed by Christensen, Jorgenson, and Lau (1971, 1973), the generalized Leontief function introduced by Diewert (1971), and the generalized Cobb-Douglas function due to Diewert (1973a). These functional forms are proposed mainly to deal with the single output technology, but they can be easily adapted to the multiple output technology. The multiproduct functional forms are general in that they can deal with multiple outputs, variable elasticities of substitution among inputs, variable elasticities of transformation among outputs, and variable scale economies. As such, they offer substantial gains in flexibility, and enable us to test important hypotheses

concerning the underlying structure of production.

In general, the proposed flexible forms are a member of a class of functions known as general quadratic flexible forms (Blackorby, Primont, and Russell 1977):

$$F(q) = \sum_{i \in I} q_i f_i(q_i) + \sum_{i \in I} \sum_{j \in I} b_{ij} f_i(q_i) f_j(q_j) \quad (3.9)$$

where a_i and b_{ij} are the fixed parameters of the first-and second-order terms with respect to an arbitrary argument $f(q)$, and $b_{ij} = b_{ji}$ are imposed as a symmetry constraint. We adopt the translog form in this research, since it is the most popular of the flexible specification for neoclassical production and cost functions. The translog function is a second-order logarithmic approximation to an unknown function. It is obtained by choosing the natural (or common) logarithm as the metric for F and for each argument f_i .

The translog specification of the multiproduct water supply cost function given in (3.5) can be written as:¹³

$$\begin{aligned} \ln C(Y, W, Z) = & a_0 + \sum_i a_i \ln Y_i + \sum_j b_j \ln W_j + \sum_k c_k \ln Z_k \\ & + 1/2 \sum_i \sum_p a_{ip} \ln Y_i \ln Y_p + 1/2 \sum_j \sum_q b_{jq} \ln W_j \ln W_q \\ & + 1/2 \sum_k \sum_r c_{kr} \ln Z_k \ln Z_r + \sum_i \sum_j d_{ij} \ln Y_i \ln W_j \\ & + \sum_i \sum_k e_{ik} \ln Y_i \ln Z_k + \sum_j \sum_k f_{jk} \ln W_j \ln Z_k \end{aligned} \quad (3.10)$$

¹³ A complete specification of the translog cost function incorporating some constraints for cost minimization to hold is discussed in the next chapter.

where a_0 is a neutral parameter; $a_i, b_j, c_k, a_{ip}, b_{jq}, c_{kr}, d_{ij}, e_{ik},$ and f_{jk} are the technologically determined cost parameters of the first- and second-order terms with respect to outputs Y , input prices W , and operating variables Z . Additionally, $a_{ip} = a_{pi}, b_{jq} = b_{qj},$ and $c_{kr} = c_{rk}$ but $d_{ij} \neq d_{ji}, e_{ik} \neq e_{ki},$ and $f_{jk} \neq f_{kj}$ are imposed as symmetry conditions. The symmetry constraints are imposed by the equality of cross-partial derivatives of the cost function with respect to $Y, W,$ and Z , as implied by the second-order approximation property of the translog cost function. Subscripts i and p refer to outputs, i.e., residential (Y_R) and nonresidential (Y_N) water, j and q refer to prices of labor (W_L), capital (W_K), and energy (W_E), and k and r refer to capacity utilization (Z_U) and service distance (Z_D).

As stated earlier, the cost-minimizing input demand functions are derived by Shephard's Lemma. In logarithmic form, Shephard's Lemma can be written as:

$$\frac{\partial \ln C}{\partial \ln W_j} = \frac{\partial C}{\partial W_j} \frac{W_j}{C} = S_j \quad (j = L, K, E) \quad (3.11)$$

where $S_j = W_j X_j / C$, the share of the total costs accruing to input j . The sum of the S_j is constrained to be unity, as implied by the linear homogeneity in input prices.¹⁴ Once the input share equations ($S_j, j = L, K, E$) are found, they can be transformed into the input demand equations simply by:

¹⁴ Derivation of linear homogeneity in input prices is discussed in Section 3.3.

$$X_j = S_j \frac{C}{W_j} \quad (j = L, K, E) \quad (3.12)$$

By logarithmic differentiation of the translog cost function, the input share equations take the following form:

$$S_j = b_j + \sum_q b_{jq} \ln W_q + \sum_i d_{ij} \ln Y_i + \sum_k f_{jk} \ln Z_k \quad (3.13)$$

Note that the relative input shares depend upon input prices, outputs, and operating variables of the firm. The elasticity of the cost share with respect to a change in input price can be written as:

$$\frac{\partial \ln S_j}{\partial \ln W_q} = \frac{b_{jq}}{S_j} \quad (j, q = L, K, E) \quad (3.14)$$

which varies with the observed values of the cost shares.

The translog cost function has to satisfy certain regularity conditions in order to correspond to well-behaved production structures. The minimum requirements are that the translog cost function be positively linear homogeneous, monotonically increasing, and concave in input prices W . As it is very difficult to impose these requirements directly on the estimating equations, the typical procedure has been to estimate parameters which have not been conditioned on these conditions and observe whether or not these conditions are captured in the data.

Linear homogeneity in input prices is a precondition for the existence of a duality relation between the cost function and the transformation function, so it is usually adopted as a maintained

hypothesis.¹⁵ The monotonicity condition is satisfied if the cost function is increasing in input prices for given Y and Z. For the translog cost function, this condition is met, provided the fitted cost shares are positive, since input prices and costs are always positive.

Concavity of the cost function will be obtained provided the Hessian matrix of second partial derivatives with respect to input prices is negative semidefinite. For the translog cost function, the elements of the Hessian matrix are:¹⁶

$$\begin{aligned}\frac{\partial^2 C}{\partial W_j^2} &= C_{jj} = \frac{C}{W_j^2} \left(\frac{\partial \ln C}{\partial \ln W_j} \left(\frac{\partial \ln C}{\partial \ln W_j} - 1 \right) + \frac{\partial^2 \ln C}{\partial \ln W_j^2} \right) \\ &= \frac{C}{W_j^2} (S_j(S_j - 1) + b_{jj})\end{aligned}\quad (3.16a)$$

$$\begin{aligned}\frac{\partial^2 C}{\partial W_j \partial W_q} &= C_{jq} = \frac{C}{W_j W_q} \left(\frac{\partial \ln C}{\partial \ln W_j} \frac{\partial \ln C}{\partial \ln W_q} + \frac{\partial^2 \ln C}{\partial \ln W_j \partial \ln W_q} \right) \\ &= \frac{C}{W_j W_q} (S_j S_q + b_{jq})\end{aligned}\quad (3.16b)$$

¹⁵ Throughout this research, the term "hypothesis" is used to mean statistical hypothesis rather than economic hypothesis which attempts to test predictions of a theory with evidence.

¹⁶ A simple way to derive the Hessian matrix is to use (3.12) and note:

$$\frac{\partial^2 C}{\partial W_j^2} = \frac{\partial X_j}{\partial W_j} = \frac{\partial}{\partial W_j} \left(S_j \frac{C}{W_j} \right) \quad (3.15a)$$

$$\frac{\partial^2 C}{\partial W_j \partial W_q} = \frac{\partial X_j}{\partial W_q} = \frac{\partial}{\partial W_q} \left(S_j \frac{C}{W_j} \right) \quad (3.15b)$$

where b_{jj} and b_{jq} are the coefficient of the own and cross product terms of the input prices from the estimated translog cost function. Since the cost function is linear homogeneous in input prices, the Hessian matrix will be singular. However, concavity will be checked by calculating the principal minors or eigenvalues of the Hessian matrix.¹⁷

As previously stated, the translog cost function is not an exact representation but a second order approximation to the underlying unknown true function.¹⁸ To view the translog cost function as a second order approximation, we can replace $\ln Y_i$ by $(\ln Y_i - \ln Y_i^0)$, $\ln W_j$ by $(\ln W_j - \ln W_j^0)$, and $\ln Z_k$ by $(\ln Z_k - \ln Z_k^0)$. Then the point (Y_i^0, W_j^0, Z_k^0) is the point of approximation or expansion. The translog cost function as written in (3.10) can be viewed as an approximation around the unit point, so that $(Y^0, W^0, Z^0) = (1, 1, 1)$ and $\ln(Y^0, W^0, Z^0) = (0, 0, 0)$.

However, approximation around the unit point may result in poor approximation of the true function if the actual values of the variables

¹⁷ Lau (1974, 1978) proposes the necessary and sufficient inequality restrictions for monotonicity and concavity which could be imposed on the estimating equation and which requires the use of nonlinear estimation. Humphrey (1981) proposes an alternative, mixed estimation procedure which utilizes prior stochastic information on necessary second-order conditions for the cost function concavity. Diewert and Parkan (1979) extend the linear programming tests suggested by Afriat (1972) and Hanoch and Rothschild (1972) of regularity conditions on production functions and the related tests for the consistency of a data set with the hypothesis of production efficiency or with the hypotheses of cost minimization or profit maximizing behavior; see also Burley (1980).

¹⁸ The translog functional form is called "flexible" since it is capable of providing a second-order approximation to an arbitrary function. Two concepts of approximation, both satisfied by the translog function, have been distinguished by Lau (1974). In the first concept, function $H(Q)$ provides a second-order differential approximation to $G(Q)$ at Q_0 provided $H(Q_0) = G(Q_0)$, and

are far removed from the unit point. Therefore, instead of using the unit point as the point of approximation, it may be preferable to use either the sample mean or the current value of the variable as the point of approximation. The use of the sample means as the point of approximation amounts to normalizing each variable to have a mean of one and standard deviation of zero, since, for example, $(\ln Y_i - \ln Y_i^0)$ equals $\ln (Y_i/Y_i^0)$. Then, a number of important economic calculations are simplified at this point of approximation, which corresponds to a hypothetical firm whose values of input prices, outputs, and operating variables (capacity utilization and service distance) are those of the sample arithmetic means.

Since the translog cost function is an approximation function, a problem which arises is the accuracy of the approximation. If the

$$\begin{aligned} \left. \frac{\partial H}{\partial Q} \right|_{Q=Q_0} &= \left. \frac{\partial G}{\partial Q} \right|_{Q=Q_0} \\ \left[\frac{\partial^2 H}{\partial Q_i \partial Q_j} \right]_{Q=Q_0} &= \left[\frac{\partial^2 G}{\partial Q_i \partial Q_j} \right]_{Q=Q_0} \end{aligned}$$

In the second concept, $H(Q)$ provides a second-order numerical approximation to $G(Q)$ at Q_0 if $H(Q_0) = G(Q_0)$, and

$$|H(Q) - G(Q)| \leq \frac{K \|Q - Q_0\|^3}{\|Q_0\|^3}$$

for all Q in a given neighborhood of Q_0 , where K is a constant depending on Q_0 and the chosen neighborhood, and $\|Q\|$ denotes a norm of Q . Any function $H(Q)$ that can be interpreted as a second-order Taylor's series approximation to $G(Q)$ around Q_0 is both a numerical and a differential second-order approximation. A translog function can be interpreted in the same fashion.

translog function is calibrated to provide a second order approximation at a point--in this case, mean of the (Y, W, Z) vector, then approximation of this order holds only in a small neighborhood of the point. Further, the basic properties of the true function such as monotonicity and concavity are inherited by the approximating function only in this neighborhood. In other regions of interest, the form may be in poor approximation to the true function, and may even fail to satisfy globally the basic properties of the true function (the translog function does not satisfy the regularity conditions globally unless it collapses into a Cobb-Douglas form). Moreover, the qualitative implications of the calibrated approximation may depend on the point of approximation. For example, the economic effects of interest in comparative statics can be implausibly related to different points in the domain of the functional expansion.¹⁹

This brings up the problem as to how to interpret violations of essential conditions at points other than the point of approximation. In some sense, a "good" approximation will not violate essential conditions within the sample, but the more important question is whether satisfaction of the essential conditions throughout the sample range (or possibly excepting fringe points) is indicative of a satisfactory approximation. Wales (1977) indicates that translog estimates of important elasticities are quite good for cases in which monotonicity and concavity hold for all points in a sample, and generally very good if the

¹⁹ See Denny and Fuss (1975) for an approximation interpretation of the translog function and implications for separability and aggregation.

violations occur only for a small number of points; see also Christensen and Caves (1980). Thus, the occasional violation of regularity conditions does not preclude obtaining good parameter estimates, nor does it necessarily undermine the assumption of cost minimization. Rather, it implies that the translog function is a poor approximation to its true functional representation.

3.3 TESTS OF THE UNDERLYING PRODUCTION STRUCTURE OF WATER SUPPLY

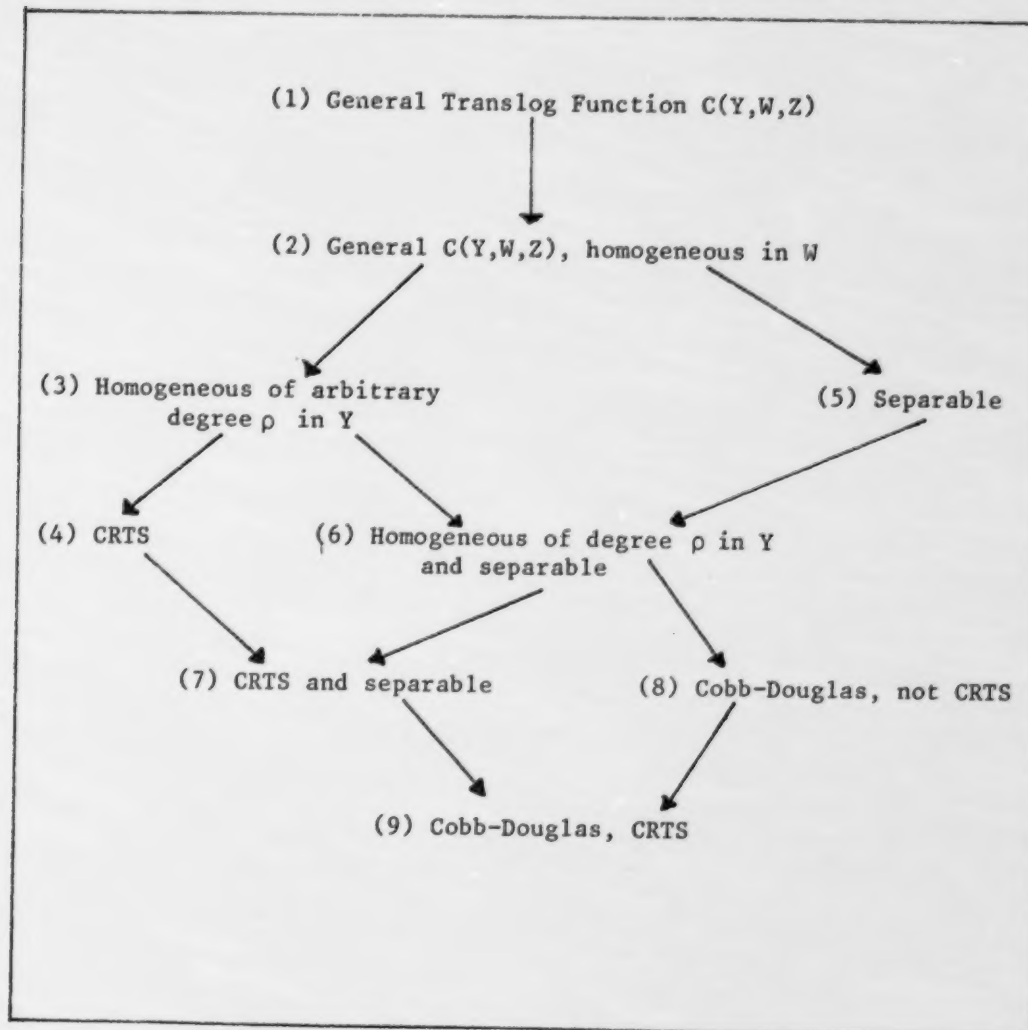
The translog multiproduct cost function (3.10) is entirely general and does not impose any restrictions on the underlying structure of the water supply technology. Nevertheless, using this translog cost function, we can test the presence of the various specialized production structures of water supply with respect to homogeneity, separability, and homotheticity as well as nonjointness in production.²⁰ Since these specialized production structures are restrictions imposed on the original translog multiproduct cost function, they are a series of nested models for the same family of parametric cost functions. As indicated in the previous chapter, previous studies on water supply costs employ very restrictive assumptions about the technology of water supply as well as very restrictive functional forms. Thus, tests of the above restrictions may provide some clues as to the validity of the existing methodology of specifying the structure of water supply.

Figure 2 illustrates the structure of the nested models concerning

²⁰See Christensen and Greene (1976) and Corbo and Meller (1979) for the single product case.

Figure 2

Tests of Underlying Production Structure Using Translog Cost Functions



homogeneity and separability as well as restricted functional forms such as Cobb-Douglas. Beginning with a general translog cost function (1) and imposing the restrictions necessary for cost minimization (2), then the restrictions necessary for production to be homogeneous of an arbitrary degree in outputs (3), and for production to be subjected to constant returns to scale (CRTS) in outputs (4) can be imposed. The restrictions necessary for production to be separable (5), or separable and homogeneous in outputs (6) can also be examined. It is possible to evaluate the combinations of restrictions needed if production is subject to CRTS and is separable (7); if production can be described by a Cobb-Douglas technology of arbitrary degree in Y (8); and if production can be described by a CRTS, Cobb-Douglas technology (9).

Homogeneity in Input Prices

The basic theory of cost minimization requires that the cost function be homogeneous of degree one in input prices W . Homogeneity of degree one in W for given Y and Z requires:

$$C(Y, tW, Z) = tC(Y, W, Z) \quad (3.17)$$

for all scalars $t \geq 0$.

For the translog cost function,

$$\ln C(Y, tW, Z) = \ln(tC(Y, W, Z)) = \ln t + \ln C(Y, W, Z) \quad (3.18)$$

This produces the following parameter restrictions:²¹

²¹ See the Appendix at the end of this chapter for the derivation of parameter restrictions necessary for homogeneity in input prices and outputs.

$$\begin{aligned}
\sum_j b_j &= 1 \quad (j = L, K, E) \\
\sum_j b_{jq} &= \sum_q b_{jq} = \sum_j \sum_q b_{jq} = 0 \quad (j, q = L, K, E) \\
\sum_j d_{ij} &= 0 \quad (j = L, K, E; i = R, N) \\
\sum_j f_{jk} &= 0 \quad (j = L, K, E; k = U, D)
\end{aligned} \tag{3.19}$$

Homogeneity in Outputs

A transformation function is homogeneous of degree ρ in outputs Y , if when all inputs are scaled by the same proportion, all outputs are scaled by the same proportion; see Lau (1972), and Brown, Caves, and Christensen (1980). Homogeneity of the structure of production requires that the multiproduct cost function be homogeneous in outputs. Since $C(Y, W, Z)$ is symmetric in W and Y , the above result on homogeneity in input prices can be extended to derive the necessary restrictions for homogeneity in outputs Y of an arbitrary degree. For the translog cost function, homogeneity of degree ρ in Y requires that:

$$\begin{aligned}
\sum_i a_i &= \rho \quad (i = R, N) \\
\sum_i a_{ip} &= \sum_p a_{ip} = \sum_i \sum_p a_{ip} = 0 \quad (i, p = R, N) \\
\sum_i d_{ij} &= 0 \quad (i = R, N; j = L, K, E) \\
\sum_i e_{ik} &= 0 \quad (i = R, N; k = U, D)
\end{aligned} \tag{3.20}$$

Corresponding to the multiproduct cost function which is homogeneous of degree ρ in outputs, the transformation function is homogeneous

of degree $1/\rho$ in output. Constant returns to scale (CRTS) results if the cost function is homogeneous of degree one in outputs, i.e., $\rho = 1$.

To test for homogeneity of an arbitrary degree ρ in Y , the cost function can be estimated subject to the constraints given by (3.19) and (3.20). The results can be compared with those on the unconstrained cost function (3.10) to see if the two equations are "significantly" different.²² To test for CRTS, the additional constraint that $\sum_i a_i = 1$ is imposed and the results must be compared with the equation without the imposed constraints.

Input-Output Separability²³

A common procedure in the study of multiproduct firms is the use of a single variable to represent output, no matter how diverse are its actual components. Formally, this is equivalent to assuming a separable transformation frontier of the form:

$$T(Y, X, Z) = T(G(Y), X, Z) \quad (3.21)$$

This requires the separability of outputs from inputs such that the transformation function can be expressed as:

$$G(Y, Z) = H(X, Z) \quad (3.22)$$

The test for separability is then equivalent to the test for the existence of an output aggregate.

²² The precise testing procedure will be discussed in the next chapter.

²³ Input-output separability is also referred to as homothecity for the single product technology. But this is different from the separability of labor, capital, and energy from other inputs stated earlier in Section 3.1.

The multiproduct cost function associated with the input-output separability can be written as (Hall 1972):

$$C(Y, W, Z) = C(G(Y), W, Z) \quad (3.23)$$

Separability then implies that the relative marginal costs of any two outputs (or the marginal rates of transformation between two outputs), residential and nonresidential water, depend only on the composition of the output mix and are independent of input prices (or the composition of inputs). Separability for the translog cost model requires:

$$\frac{\partial}{\partial W_q} \left(\frac{\partial \ln C}{\partial \ln Y_R} \right) / \left(\frac{\partial \ln C}{\partial \ln Y_N} \right) = 0 \quad (q = L, K, E) \quad (3.24)$$

This implies that:

$$\frac{\partial}{\partial W_q} \left(\frac{a_R + \sum_p a_{Rp} \ln Y_p + \sum_j d_{Rj} \ln W_j + \sum_k e_{Rk} \ln Z_k}{a_N + \sum_p a_{Np} \ln Y_p + \sum_j d_{Nj} \ln W_j + \sum_k e_{Nk} \ln Z_k} \right) = 0 \quad (3.25)$$

At the point of approximation, this implies that:

$$\begin{aligned} a_{R Nq}^d &= a_{N Rq}^d \\ d_{Rq} a_{Nj} &= d_{Nq} a_{Rj} \end{aligned} \quad (3.26)$$

A sufficient condition to satisfy these restrictions is that:²⁴

²⁴ An easy way to check this condition is to rewrite the separable cost function in the form:

$$C(Y, W, Z) = G(Y, Z) H(W, Z) \quad (3.27a)$$

or in logarithmic form

$$\ln C(Y, W, Z) = \ln G(Y, Z) + H(W, Z) \quad (3.27b)$$

then $d_{ij} = 0$ is the sufficient condition for separability.

$$d_{ij} = 0 \quad (i = R, N; j = L, K, E) \quad (3.28)$$

That is, all interaction terms between Y_i and W_j must be zero if the transformation process is input-output separable.

If the separability and homogeneity of an arbitrary degree in Y is assumed, the further restrictions given in (3.28) and (3.20) must be added.

One of the assumptions in the specification of the cost function for water supply was that the water utility provides multiple-- and distinct-- outputs to residential and nonresidential customers. The validity of this assumption can be tested statistically by the input-output separability. To test for separability, the translog cost function subject to the restrictions given in (3.28) must be estimated and the results compared with the cost function without the restrictions.

The results obtained so far can be summarized beginning with (1) the translog cost function given in (3.10). Then the restriction in (3.19) is imposed, which subsequently assumes to hold for cost minimization. Thus, the ensuing analysis assumes that restrictions (3.19) hold, as well as those needed for homogeneity and separability. Homogeneity of an arbitrary degree in Y (3) requires that (3.20) hold, while constant returns to scale (CRTS) (4) imposes the additional restriction that $\sum_i a_i = 1$. Separability (5) requires that (3.28) holds. The restrictions implied by (3.19) and (3.28) can then be combined to test for separability and homogeneity (6). By further requiring that (3.20) and (3.28) hold and that $\sum_i a_i = 1$, separability and CRTS (7) is testable.

Moreover, by requiring that (3.20) and (3.28) hold and that $b_{jq} = 0$,²⁵ a test for Cobb-Douglas technology, with no CRTS restriction (8) can be performed. Finally, by requiring that (3.20) and (3.28) hold and that $b_{jq} = 0$ and $\sum_i a_i = 1$, a test for CRTS, Cobb-Douglas technology (9) can be conducted.

Thus, the Cobb-Douglas specification of technology is of a restrictive form and is a special case of the translog specification. Conveniently, the Cobb-Douglas cost function is obtained by requiring all the second-order terms of the translog cost function to vanish. Using the above restrictions, however, it is possible to test the appropriateness of the restrictive functional forms such as Cobb-Douglas which are widely used in the existing studies.

Nonjoint Production and Economies of Joint Production²⁶

When production is joint, the production of delivered water to residential and nonresidential customers uses joint facilities in treatment and distribution; hence joint costs will be incurred in the provision of water to different customer classes. When production is nonjoint, each output is produced by its own distinct production process; see Samuelson (1966), Lau (1972), and Kohli (1982). Therefore, the joint facilities are not available for use in the provision of water

²⁵ The condition that $b_{jq} = 0$ is equivalent to unitary elasticities of substitution among inputs. A general discussion of elasticities of substitution is presented in the next section on input demand.

²⁶ Here "jointness" and "nonjointness" is used in the production sense, encompassing both common and joint costs as used in the water supply literature.

to residential and nonresidential users and joint costs do not exist. As a result, the nonjoint cost function can be written as the sum of two cost functions, each of which is a function of a separate output. Then,

$$\begin{aligned} & C(Y_R, Y_N; W_L, W_K, W_E; Z_U, Z_D) \\ &= C^R(Y_R; W_L, W_K, W_E; Z_U, Z_D) + C^N(Y_N; W_L, W_K, W_E; Z_U, Z_D) \end{aligned} \quad (3.29)$$

Thus, we have two separate, residential and nonresidential, cost functions and each cost function has input prices and operating variables (capacity utilization and service distance) as the same arguments. However, other than input prices and operating variables, the residential water cost function is a function only of residential water, and the nonresidential water cost function is a function only of nonresidential water.

The nonjointness condition implies that:

$$\frac{\partial^2 C}{\partial Y_R \partial Y_N} = 0 \text{ for } R \neq N \quad (3.30)$$

Consequently, if we can relate the coefficients of the translog cost function to the second derivatives of the underlying cost function and impose the condition that $\partial^2 C / \partial Y_R \partial Y_N = 0$, we can then determine the restrictions on the coefficients of the translog function to ensure that production is nonjoint.

For the translog cost model,

$$\begin{aligned}
\frac{\partial^2 C}{\partial Y_N \partial Y_N} &= \frac{C}{Y_R Y_N} \left(\frac{\partial^2 \ln C}{\partial \ln Y_R \partial \ln Y_N} + \frac{\partial \ln C}{\partial \ln Y_R} \frac{\partial \ln C}{\partial \ln Y_N} \right) \\
&= \frac{C}{Y_R Y_N} (a_{RN} + \frac{\partial \ln C}{\partial \ln Y_R} \frac{\partial \ln C}{\partial \ln Y_N}) \quad (3.31)
\end{aligned}$$

The nonjointness restrictions are therefore:

$$a_{RN} + a_R a_N = 0 \quad (3.32)$$

when the data have been scaled so that all $Y_i = W_j = Z_k = 1$ at the point of approximation and that $\partial \ln C / \partial \ln Y_R = a_R$ and $\partial \ln C / \partial \ln Y_N = a_N$.

Nonjointness in production implies that there are no economies or diseconomies of jointness in producing the outputs. If there are economies of joint production, we will have the condition:

$$C(Y, W, Z) < \sum_i C^i(Y_i, W, Z) \quad (3.33)$$

That is, the cost of producing residential and nonresidential water jointly is less than the sum of the costs of producing separately. If there are diseconomies of jointness in production, we will have:

$$C(Y, W, Z) > \sum_i C^i(Y_i, W, Z) \quad (3.34)$$

That is, the cost of producing outputs jointly exceeds the costs of producing separately. Obviously, it is more efficient to provide residential and nonresidential water supply jointly only if there are economies of jointness, because it is always cheaper to provide two water supplies jointly than separately.

Economies (or diseconomies) of jointness in production can be

tested in the translog framework.²⁷ A sufficient condition for the cost function to display economies of joint production is that:

$$\frac{\partial^2 C}{\partial Y_R \partial Y_N} < 0 \quad (3.35)$$

That is, the production of residential and nonresidential water supply exhibits cost complementarity. At the point of approximation, the complementary production relation is obtained if (3.32) is negative, that is, if:

$$a_{RN} + a_R a_N < 0 \quad (3.36)$$

As indicated before, when nonjoint costs exists, it is possible to measure the cost of residential and nonresidential water supply separately. This is a restrictive assumption adopted in the existing studies in water supply, particularly in the studies of water rate design. As will be shown later, this restrictive specification has implications for the measurement of scale economies and other production characteristics of water supply. However, using the nonlinear restrictions on the interaction coefficients of the output variables as given in (3.32), it is possible to test for nonjoint production and to check the validity of the existing specification.

²⁷ For a more general discussion of economies of joint production under the name of "economies of scope," see Panzar and Willig (1981), and Baumol and Braunstein (1977), Willig (1979), and Baumol, Panzar, and Willig (1982). For a survey of empirical analyses of multiproduct firms and industries and of economies of scope, see Bailey and Friedlaender (1982), and Baumol, Panzar, and Willig (1982, Chapter 15).

3.4 STRUCTURE OF INPUT DEMAND

Given the general cost function, the relevant input demand functions are a transformation of this form. Demand for each input depends not only on the price of its own input but also on the price of other inputs, the outputs, and the operating variables of the firm. This raises the issue of how much each input varies in response to changes in input prices, outputs, and operating variables.

Two of the most common summary measures of price responsiveness of input demands are the Allen (partial) elasticities of substitution²⁸ and the price elasticities of input demands. Estimates of the two elasticities are calculated from the estimated translog cost function, under the assumption that outputs are held fixed but that the quantities of all other inputs are permitted to vary. A direct way to derive the price elasticities of input demands and the elasticities of substitution is to make use of the input demand equations (.12).

The own price elasticity of demand for input j is defined as:

$$\eta_{jj} = \frac{\partial \ln X_j}{\partial \ln W_j} = \frac{\partial \ln(S_j C/W_j)}{\partial \ln W_j} \quad (3.37)$$

²⁸There are at least two other concepts of elasticities of substitution that deal with more than two inputs: (1) the Morishima (or direct) elasticity of substitution (Blackorby and Russell 1981) which is a direct extension of the traditional elasticity of substitution and (2) the Hicks (or shadow) elasticity of substitution (McFadden 1963) which measures the percent change in the ratio of input prices in response to an exogenous percentage change in input quantity ratio. For more on these elasticities of substitution, see Mundlak (1969), Sato and Koizumi (1973), and Kang and Brown (1981).

which reduces to:

$$\eta_{jj} = \frac{\partial \ln S_j}{\partial \ln W_j} + S_j - 1 \quad (3.38)$$

After finding an expression for $\partial \ln S_j / \partial \ln W_j$ from (3.13), the own price elasticity of input demand can be written as:

$$\eta_{jj} = \frac{1}{S_j} (b_{jj} + S_j (S_j - 1)) \quad (3.39)$$

The cross price elasticity of demand for input j with respect to the price of input q is defined as:

$$\eta_{jq} = \frac{\partial \ln X_j}{\partial \ln W_q} = \frac{\partial \ln (S_j C / W_j)}{\partial \ln W_q} \quad (3.40)$$

Since $\partial \ln W_j / \partial \ln W_q = 0$, this reduces to:

$$\eta_{jq} = \frac{\partial \ln S_j}{\partial \ln W_q} + S_q \quad (3.41)$$

After finding an expression for $\partial \ln S_j / \partial \ln W_q$ from (3.13), we get:

$$\eta_{jq} = \frac{1}{S_j} (b_{jq} + S_j S_q) \quad (3.42)$$

Stable input demand functions require that each own price elasticity η_{jj} be nonpositive. Moreover, the derivative of the own

price elasticity with respect to the cost share is:

$$\frac{\partial \eta_{jj}}{\partial s_j} = 1 - \frac{b_{jj}}{s_j^2} \quad (3.43)$$

Since b_{jj} is a constant, the elasticity will fall as the share increases and increase as the share falls--a property of any translog cost function.

The Allen elasticities of substitution are analytically related to the own and cross price elasticities of input demand in the following way (Allen 1938, Uzawa 1962):

$$\sigma_{jj} = \frac{1}{s_j} \eta_{jj} = \frac{1}{s_j^2} (b_{jj} + s_j(s_j - 1)) \quad (3.44a)$$

$$\sigma_{jq} = \frac{1}{s_q} \eta_{jq} = \frac{1}{s_j s_q} (b_{jq} + s_j s_q) \quad (3.44b)$$

Thus, the Allen elasticities of substitution (σ_{jq}) can be interpreted as normalized price elasticities, where the normalization is chosen such that the elasticities of substitution are invariant to the ordering of the inputs. Accordingly, $\sigma_{jq} = \sigma_{qj}$, as guaranteed by the translog symmetry condition $b_{jq} = b_{qj}$, although, in general, $\eta_{jq} \neq \eta_{qj}$.

A positive elasticity of substitution implies that inputs are substitutes, while a negative elasticity of substitution implies that inputs are complements. A special case of a unitary elasticity of substitution between two inputs, i.e., a Cobb-Douglas technology, holds if $\sigma_{jq} = 1$. This amounts to testing $b_{jq} = 0$ in the translog cost func-

tion. If $\sigma_{jq} = 0$, input j and q are independent which corresponds to a Leontief technology of fixed coefficient production. Notice that elasticities of substitution and price elasticities of input demand are functions of the cost shares; hence they vary across the firm in the sample.

The question of whether or not there are substitution possibilities among water supply inputs in water utilities has an important bearing upon the policy maker. Consider the effect of higher priced energy on the investment behavior of the utility manager, a situation well described by Berndt and Wood (1975) in their study on energy:

If energy and capital are substitutable, *ceteris paribus*, then higher priced energy will increase the demand for a new capital good. If energy and capital are complementary, however, then *ceteris paribus*, higher priced energy will dampen the demand for energy and the demand for new plant and equipment. More generally, if it were found that possibilities for substitution between energy and nonenergy inputs are extremely limited, then we might expect that the adjustments by industry to higher priced energy will be somewhat difficult, that unit cost may rise considerably, that the composition of output may shift away from energy intensive products, and that significant changes in the underlying technological structure may be required.

It is useful to calculate the elasticity of input demand with respect to changes in output, i.e., residential and nonresidential water usage, holding input prices and operating variables constant. Utilizing (3.12), this elasticity can be written as:

$$\frac{\partial \ln X_j}{\partial \ln Y_i} = \frac{\partial \ln (C_j C / W_j)}{\partial \ln Y_i} \quad (3.45)$$

Since $\partial \ln W_j / \partial \ln Y_i = 0$, this reduce to:

$$\frac{\partial \ln X_j}{\partial \ln Y_i} = \frac{\partial \ln S_j}{\partial \ln Y_i} + E_{CY_i} \quad (3.46)$$

where $E_{CY_i} = \partial \ln C / \partial \ln Y_i$, the elasticity of cost with respect to the i th output. The j th input is said to be superior if the output elasticity of that input is greater than 1; it is normal if its output elasticity is less than 1 (but greater than 0); and it is inferior if the elasticity is less than 0.

For the translog cost function, substituting $\partial \ln S_j / \partial \ln Y_i = d_{ij}/S_j$ and deriving an expression for E_{CY_i} ,²⁹ we obtain:

$$\begin{aligned} \frac{\partial \ln X_j}{\partial \ln Y_i} = & d_{ij}/S_j + a_i + \sum_p a_{ip} \ln Y_p + \sum_j d_{ij} \ln W_j \\ & + \sum_k e_{ik} \ln Z_k \end{aligned} \quad (3.47)$$

Thus, the elasticity of the demand for labor with respect to a change in residential demand for water can be written as:

$$\begin{aligned} \frac{\partial \ln X_L}{\partial \ln Y_R} = & d_{LR}/S_L + a_R + a_{RR} \ln Y_R + a_{RN} \ln Y_N + d_{RL} \ln W_L \\ & + d_{RK} \ln W_K + d_{RE} \ln W_E + e_{RU} \ln Z_U + e_{RD} \ln Z_D \end{aligned} \quad (3.48)$$

a_{ip} 's, d_{ij} 's, and e_{ik} 's reflect the influences of outputs, input prices, and operating variables, respectively, upon the output elasticity of input demand. Therefore, given parameter estimates from the translog cost function, we can compute the output elasticity at sample

²⁹ See equation (3.52) in the discussion of scale economies.

observation of outputs, input prices, and operating variables.³⁰ For a representative firm evaluated at the approximation points of the translog function, the output elasticity can be measured by $d_{ij}/S_j + a_i$.

Changes in operating variables, i.e., capacity utilization and service distance will affect the demand for inputs and input requirements necessary for the production of delivered water. The effect of changes in capacity utilization and service demands can be computed by the use of the elasticity of input demand with respect to capacity utilization and service distance.

The elasticity of demand for input j with respect to operating variable k is defined as:

$$\frac{\partial \ln X_j}{\partial \ln Z_k} = \frac{\partial \ln S_j}{\partial \ln Z_k} + \frac{\partial \ln C}{\partial \ln Z_k} \quad (3.49)$$

where $\partial \ln C / \partial \ln Z_k$ is the cost elasticity with respect to operating variable k .³¹

For the translog cost function, we obtain:

$$\begin{aligned} \frac{\partial \ln X_j}{\partial \ln Z_k} = & d_{jk}/S_j + c_k + \sum_r d_{kr} \ln Z_r + \sum_i e_{ik} \ln Y_i \\ & + \sum_j f_{jk} \ln W_j \end{aligned} \quad (3.50)$$

³⁰ Note that constant returns to scale imply that the sum of the output demand elasticities for each input is unity. Note also that the inverse of the output elasticity can be considered as the elasticity of output with respect to inputs.

³¹ A detailed discussion of this elasticity is presented in Section 3.6.

c_{kr} 's, e_{ik} 's, and f_{jk} 's show the responses of operating variables, outputs, and input prices, respectively, upon the input elasticity of operating variables. Thus, the elasticity of the demand for capital with respect to a change in capacity utilization can be written as:

$$\frac{\partial \ln X_K}{\partial \ln Z_U} = f_{KU}/S_K + c_U + c_{UU} \ln Z_U + c_{UD} \ln Z_D + e_{RU} \ln Y_R \\ + e_{NU} \ln Y_N + f_{LU} \ln W_L + f_{KU} \ln W_K + f_{EU} \ln W_E \quad (3.51)$$

For a representative utility evaluated at the approximation points of the translog cost function, the elasticity of input demand with respect to an operating variable can be calculated by $f_{jk}/S_j + c_k$.

3.5 SCALE ECONOMIES

The question of whether and over what range there are economies or diseconomies of scale is important for both the utility's business decisions and policy analysis. Scale economies are a determinant of investment policies in water utilities, especially when this industry grows. If there are economies of scale and growing demand, utilities may find it profitable to add more capacity than they expect to use in the immediate future. Further, if there are economies of scale in water supply, large utilities could produce at lower average cost than smaller firms; thus a competitive equilibrium would not be sustainable. Then a valid justification could be provided for economic regulation of utilities as a means of ensuring competitive markets in water supply, thereby bringing about a more efficient resource allocation.

Also, as will be demonstrated later, the evidence of scale economies is central to the applicability of marginal cost pricing. Moreover, if economies of scale actually exist in water supply, then a valid argument can be made for the establishment of a large central water supply system in order to gain the benefits of these economies. It has been often suggested that economies of scale associated with "regionalization" would benefit small systems. For these reasons, previous studies on water supply have been largely concerned with estimating scale economies.

A good measure of local scale economies is the cost elasticity of output evaluated along the expansion path where input prices (and operating variables) are held constant and costs are minimized at every level of output; see Hanoch (1975a) and Brown and Chachere (1980). The cost elasticity of output is the proportional increase in cost resulting from a proportional increase in the level of output. Since there are two outputs, residential and nonresidential water supply, in the model, there are two corresponding cost elasticities of output. For the translog multiproduct cost function, the cost elasticity of the i th output can be expressed as:

$$\begin{aligned}
 E_{CY_i}(Y, W, Z) &= \frac{\partial \ln C}{\partial \ln Y_i} \quad (i = R, N) \\
 &= a_i + \sum_p a_{ip} \ln Y_p + \sum_j d_{ij} \ln W_j + \sum_k e_{ik} \ln Z_k \quad (3.52)
 \end{aligned}$$

The cost elasticity can be used to measure the output-specific scale

economies,³² i.e., scale economies for residential or nonresidential water supply.

The aggregate or overall scale economies for the multiproduct firm such as water utilities is the sum of the cost elasticities of a single output (Baumol 1975, Panzar and Willig 1977b):³³

$$E_{CY}(Y, W, Z) = \sum_i (\partial \ln C / \partial \ln Y_i) = \sum_i E_{CY_i}(Y, W, Z) \quad (3.53)$$

³² It should be noted that the cost elasticity of an individual output biases the output (product)-specific scale economies. By definition (Baumol, Panzar, and Willig 1982), scale economies specific to product i are measured by $(\partial C / \partial Y_i)(Y_i / C^i)$, where C^i is the incremental cost of producing the product i at a specific level of output as compared with not producing it at all, e.g., the incremental cost of residential water supply is:

$$C^R = C(Y_R, Y_N; W; Z) - C(0, Y_N; W; Z)$$

and $\partial C / \partial Y_i = \partial C^i / \partial Y_i$. Output-specific scale economies are related to the cost elasticity of the individual output by:

$$\frac{\partial C}{\partial Y_i} \frac{Y_i}{C^i} = \frac{\partial C}{\partial Y_i} \frac{Y_i}{C^i} \frac{C}{C} = E_{CY_i} \frac{C}{C^i}$$

Since C/C^i is always greater than one, $(\partial C / \partial Y_i)(Y_i / C^i)$ is greater than E_{CY_i} . This implies that the cost elasticity of the individual output overestimates the true output-specific scale economies. The cost elasticity of the single product equals the output-specific scale economies only if production is nonjoint; since $\partial \ln C / \partial \ln Y_i = \partial \ln C^i / \partial \ln Y_i$ in this case.

³³ Alternatively, the multiproduct scale economies can be defined as:

$$E_{CY} = \frac{\sum_i Y_i MC_i}{C(Y, W, Z)}$$

where MC_i is the marginal cost of the i th output. Thus, the degree of multiproduct scale economies equals the weighted sum of marginal costs divided by total costs, with the weights representing the different levels of output i . If marginal costs are identical, the multiproduct scale economies collapses into the conventional single-product measure of scale economies. For relationships among multiproduct scale economies, output-specific scale economies, and scope economies, see Baumol, Panzar, and Willig (1982).

Using E_{CY} as a local measure of overall scale economies, E_{CY} is less than one for increasing returns to scale (economies of scale); it is greater than one for decreasing returns to scale (diseconomies of scale); and it equals to one for constant returns to scale.

It is important to emphasize that scale economies (output-specific or overall) depend on the output levels, input prices, and operating variables that are observed as well as the parameters of the cost function. The effect of an operating variable on scale economies, for example, can be seen by:

$$\partial \ln E_{CY_i} / \partial \ln Z_k = e_{ik} / E_{CY_i} \quad (3.54)$$

This result, which follows from the flexible nature of the translog function, allows the estimate of scale economies to differ at every observation in the sample space. However, a hypothetical firm with sample mean output levels, input prices, and operating variables can be computed to determine the presence and degree of both output-specific and overall scale economies.

It should be noted that the measure of scale economies (output-specific or overall) reflects a significant nonhomothetic technology present in water supply; thus no restrictions are imposed on the structure of production. However, it is of interest to investigate the impact of the imposition of restrictions on the measure of scale economies.

When technology is input-output separable, i.e., $d_{ij} = 0$ ($i = R, N$; $j = L, K, E$), the cost elasticity of the i th output is given by:

$$ECY_i = a_i + \sum_p a_{ip} \ln Y_p + \sum_k e_{ik} \ln Z_k \quad (3.55)$$

When the production structure is homogeneous of degree ρ in output, i.e., $\sum_i a_i = \rho$, $\sum_i a_{ip} = 0$, and $\sum_i e_{ik} = 0$ then

$$ECY_i = a_i \quad (3.56a)$$

and

$$ECY = \sum_i a_i \quad (3.56b)$$

Thus, the degree of scale economies does not depend on any variables and is constant for all values of outputs, input prices, and operation variables. Unitary elasticity of substitution, i.e., $b_{jq} = 0$ ($j, q = L, K, E$) has no impact on the measurement of scale economies; hence the form of scale economies is the same as that given in (3.52).

3.6 RELATION BETWEEN OPERATING VARIABLES AND COSTS

It was observed before that changing operating characteristics of the firm such as capacity utilization and service distance have significant effects on costs and production characteristics such as scale economies and input requirements. The impact upon input demands and scale economies of changes in operating variables were discussed in the last two sections. Now, the effects upon costs of changing operating variables are analyzed.

Using the flexible specification of the technology of water supply it is possible to quantify or estimate the impact upon costs of

a change in capacity utilization or service distance for a given utility. For the translog cost function, we can calculate the elasticity of cost with respect to operating variable k :

$$\frac{\partial \ln C}{\partial \ln Z_k} = c_k + \sum c_{kr} \ln Z_r + \sum e_{ik} \ln Y_i + \sum f_{jk} \ln W_j \quad (3.57)$$

Thus, the cost elasticity with respect to capacity utilization can be written as:

$$\begin{aligned} \frac{\partial \ln C}{\partial \ln Z_U} &= c_U + c_{UU} \ln Z_U + c_{UD} \ln Z_D + e_{RU} \ln Y_R + e_{NU} \ln Y_N \\ &\quad + f_{LU} \ln W_L + f_{KU} \ln W_K + f_{EU} \ln W_E \end{aligned} \quad (3.58a)$$

The cost elasticity with respect to service distance can be written as:

$$\begin{aligned} \frac{\partial \ln C}{\partial \ln Z_D} &= c_D + c_{UU} \ln Z_U + c_{UD} \ln Z_D + e_{RU} \ln Y_R + e_{NU} \ln Y_N \\ &\quad + f_{LU} \ln W_L + f_{KU} \ln W_K + f_{EU} \ln W_E \end{aligned} \quad (3.58a)$$

Thus, c_{kr} 's, e_{ik} 's, and f_{jk} 's reflect the impacts of operation variable, outputs, and input prices, respectively, on the change in costs due to a change in capacity utilization or service distance. The cost elasticity of operating variables can be calculated at the sample values of data. At the sample means, the cost elasticity with respect to operating variable k is measured by c_k .

It is useful to calculate marginal cost of an increase in operating variable. Since

$$\frac{\partial \ln C}{\partial \ln C_k} = \frac{\partial C}{\partial Z_k} \frac{Z_k}{C} \quad (3.59)$$

it is readily apparent that the marginal costs with respect to any operating variable are given by the following expression:

$$\frac{\partial C}{\partial Z_k} = \frac{C}{Z_k} \frac{\partial \ln C}{\partial \ln Z_k} \quad (3.60)$$

The marginal cost expression (3.60) has the same arguments as those of the cost elasticity of an operating variable given in (3.57).

3.7 MARGINAL COSTS AND OUTPUT SUPPLY

In the rate design, marginal cost pricing, which equates the rate (price) to marginal costs, is strongly advocated as the appropriate rate of water service. Thus, estimates of marginal costs would provide a valuable information to utility commissions whose enabling legislation requires that rates be reasonably related to the cost of providing services. Also, from the properties of the duality principle, the transformation function of water supply can be described by the relative marginal cost ratio of residential and nonresidential water supply. Thus, marginal costs are of great importance in the analysis of water supply operations. This section provides an examination of marginal costs of water supply. References to pricing problems will be made in the next section.

For the translog multiproduct cost model, the cost elasticity of output can be used to compute marginal cost of output, residential or nonresidential water supply. The marginal cost of the i th output is

obtained by the expression:

$$MC_i(Y, W, Z) = \frac{\partial C}{\partial Y_i} = \frac{C}{Y_i} E_{CY_i} \quad (i = R, N) \quad (3.61)$$

Note that marginal cost of output varies not only with the level of its output but also with the levels of other outputs, input prices, and operating variables. Then it should be useful to determine the degree to which the marginal cost responds due to changes in the variables affecting it.

To examine the effects of changes in outputs, input prices and operating variables on the marginal cost of output, we can derive the marginal cost elasticities corresponding to each variable. These elasticities can be calculated in a manner analogous to the elasticities of input demands. To simplify derivations, we make full use of the expression for the marginal cost of outputs given in (3.61).

The marginal cost elasticity of output Y_i with respect to its output Y_i ³⁴ can be given by:

$$\frac{\partial \ln MC_i}{\partial \ln Y_i} = \frac{\partial \ln (E_{CY_i} C/Y_i)}{\partial \ln Y_i} \quad (3.62)$$

which reduces to:

$$\frac{\partial \ln MC_i}{\partial \ln Y_i} = \frac{\partial \ln E_{CY_i}}{\partial \ln Y_i} + E_{CY_i} - 1 \quad (3.63)$$

³⁴ These elasticities can be considered as the inverse elasticities of transformation (or supply).

After finding an expression for $\partial \ln E_{CY_i} / \partial \ln Y_i$ from (3.52), we obtain:

$$\frac{\partial \ln MC_i}{\partial \ln Y_i} = \frac{1}{E_{CY_i}} (a_{ii} + E_{CY_i} (E_{CY_i} - 1)) \quad (3.64)$$

The marginal cost elasticity of output Y_i with respect to output Y_p is given by:

$$\frac{\partial \ln MC_i}{\partial \ln Y_p} = \frac{\partial \ln (E_{CY_i} C/Y_i)}{\partial \ln Y_p} \quad (3.65)$$

Since $\partial \ln Y_i / \partial \ln Y_p = 0$, this reduces to:

$$\frac{\partial \ln MC_i}{\partial \ln Y_p} = \frac{\partial \ln E_{CY_i}}{\partial \ln Y_p} + E_{CY_i} \quad (3.66)$$

After finding an expression for $\partial \ln E_{CY_i} / \partial \ln Y_p$ from (3.52), we get:

$$\frac{\partial \ln MC_i}{\partial \ln Y_p} = \frac{1}{E_{CY_i}} (a_i + E_{CY_i} E_{CY_p}) \quad (3.67)$$

Similarly, we can obtain the marginal cost elasticity of output Y_i with respect to input price W_j :

$$\frac{\partial \ln MC_i}{\partial \ln W_j} = \frac{\partial \ln (E_{CY_i} C/Y_i)}{\partial \ln W_j} \quad (3.68)$$

which reduces to:

$$\frac{\partial \ln MC_i}{\partial \ln W_j} = \frac{d_{ij}}{E_{CY_i}} + S_j \quad (3.69)$$

after obtaining an expression for $\partial \ln E_{CY_i} / \partial \ln W_j$ from (3.52).

And, the marginal cost elasticity of output Y_i with respect to operating variable Z_k is given by:

$$\frac{\partial \ln MC_i}{\partial \ln Z_k} = \frac{\partial \ln (E_{CY_i} C/Y)}{\partial \ln Z_k} \quad (3.70)$$

which simplifies to:

$$\frac{\partial \ln MC_i}{\partial \ln Z_k} = \frac{e_{ik}}{E_{CY_i}} + \frac{\partial \ln C}{\partial \ln Z_k} \quad (3.71)$$

after finding an expression for $\partial \ln E_{CY_i} / \partial \ln Z_k$ from (3.52).

It is tempting to conclude that economies of scale indicate declining marginal cost or vice versa. However, the degree of scale economies does not necessarily provide us with knowledge about marginal cost of output. This can be checked by investigating the curvature of the marginal cost function:

$$\frac{\partial^2 C}{\partial Y_i^2} = \frac{C}{Y_i^2} \left(\frac{\partial^2 \ln C}{\partial \ln Y_i^2} + \frac{\partial \ln C}{\partial \ln Y_i} \left(\frac{\partial \ln C}{\partial \ln Y_i} - 1 \right) \right) \quad (3.72a)$$

$$= \frac{C}{Y_i^2} (a_{ii} + E_{CY_i} (E_{CY_i} - 1)) \quad (3.72b)$$

Then the marginal cost will decline or increase depending upon the values of a_{ii} and E_{EC_i} . Suppose there are economies of scale for the i th output such that $E_{CY_i} < 1$. Then, marginal cost of the i th output will decline only if $a_{ii} < 0$ or if $a_{ii} < E_{CY_i} (E_{CY_i} - 1)$. Since these conditions will not necessarily be met, the existence of scale economies does not imply declining marginal cost of output. Consequently, we should be warned against using any observed marginal cost as an indicator of output-specific or overall scale economies.

The marginal rate of transformation between two outputs, residential and nonresidential water supply, is equal to the ratio of the corresponding marginal costs:

$$-\frac{\partial Y_N}{\partial Y_R} = \frac{MC_R}{MC_N} \quad (3.73)$$

For the translog cost function, the ratio of marginal costs of two outputs is given by:

$$\frac{MC_R}{MC_N} = \frac{Y_N}{Y_R} \frac{\partial \ln C}{\partial \ln Y_R} / \frac{\partial \ln C}{\partial \ln Y_N} \quad (3.74)$$

If both marginal costs are positive, as rational operation requires, the slope of the transformation curve is negative and the marginal rate of transformation is positive. The slope of the transformation curve is negative, since an increase in the output of residential water must result in a decrease in the output of nonresidential water.

To examine whether there are increasing opportunity costs associated with greater specialization of one output relative to other, we can

check whether the transformation function is concave or convex. This can be checked by investigating the curvature of the transformation curve by holding cost, input prices, and operating variables fixed and assessing tradeoff in the levels of residential and nonresidential water supply:

$$\frac{\partial(\partial Y_N / \partial Y_R)}{\partial Y_R} = - (C_{NR}^2 - 2C_{RN}C_{NN} + C_{NN}^2) / C_N^3 \quad (3.75a)$$

$$\partial C = \partial W_L = \partial W_K = \partial W_E = \partial Z_U = \partial Z_D = 0 \quad (3.75b)$$

where $C_i = \partial C / \partial Y_i = MC_i$, $C_{ii} = \partial^2 C / \partial Y_i^2$ given in (3.72), $C_{RN} = \partial^2 C / \partial Y_R \partial Y_N$ given in (3.31), etc.

It is normally assumed that (3.75) is negative, i.e., that the marginal rate of transformation increases as a movement is made from left to right along the transformation curve. We can evaluate this expression at the observed levels of cost, input prices, and operating variables.

The full competitive equilibrium for the multiproduct water supply firm requires that the ratio of marginal costs of two outputs, residential and nonresidential water, be equal to their respective price ratio or the slope of an isorevenue line, $R = P_R Y_R + P_N Y_N$, where R is total revenue and P_R and P_N are the prices of residential and nonresidential uses of water, respectively.

$$\frac{MC_R}{MC_N} = \frac{P_R}{P_N} \quad (3.76)$$

As indicated before, the feasibility of marginal cost pricing will be discussed in detail in the next section.

3.8 SECOND BEST PRICING RULE

It has been assumed that water utilities supply water demanded at the prices set by regulators. But, no discussion has been presented on how the prices of water are determined in the regulatory process. In their rate design, utility commissions and policy makers must consider how to price water services so as to maximize social welfare. At the same time, the enabling legislation of regulators requires that rates be reasonably related to the costs of providing services. Actual pricing may substantially differ from the optimal pricing structure. In fact, existing practices of water rates are generally carried out on the basis of average historical or embedded costs. Unfortunately, the pricing based on average costs does not maximize social welfare. Further, when there are costs of production which are shared by two or more services in the production process, it is impossible to identify costs with individual services in an unambiguous manner.

Marginal cost pricing has been advocated as an appropriate rate structure. However, water utilities often have a technology characterized by economies of scale. The supply of water involves large fixed costs and rather small marginal costs (in the sense of the costs of serving another customer rather than the costs of providing another unit of physical product). Then, as will be demonstrated below, marginal cost pricing would undoubtedly result in a negative level of profits. Hence some deviation of price from marginal cost is needed if the utility experiencing economies of scale is to break even.

In this section, we discuss some problems associated with marginal

cost pricing and provide an optimal pricing rule for water services. The pricing rule addressed in this section represents a natural step in utility regulation and may provide practical assistance to utility commissions and policy makers in their rate determination.

Suppose that marginal cost pricing is adopted so that the firm charges all outputs at marginal cost. Then we can rewrite the cost elasticity of a single output as:

$$E_{CY_i} = \frac{\partial C}{\partial Y_i} \frac{Y_i}{C} = \frac{P_i Y_i}{C} \quad (3.77)$$

where P_i is the price of the i th output ($i = R, N$) and $P_i = \partial C / \partial Y_i$, as required by marginal cost pricing. Since overall scale economies for the multiproduct water supply firm equals the sum of scale economies for individual outputs, we can write:

$$E_{CY} = \sum_i E_{CY_i} = \sum_i P_i Y_i / C = R / C \quad (3.78)$$

Overall scale economies can then be interpreted as the ratio of total revenue ($R = \sum_i P_i Y_i$) to total cost ($C = \sum_j W_j X_j$).

Economies of scale exist if and only if the revenue yielded from marginal cost pricing fall short of total costs of production. There exist diseconomies of scale if and only if total revenues exceed total costs, and there exist constant returns to scale if and only if total revenues exactly cover production costs. One implication of this result is that the firm with economies of scale cannot recover costs with marginal cost pricing. As such, marginal cost pricing alone will not ensure the solvency of the firm. The difficulties of implementing mar-

ginal cost pricing are well documented in the literature; see Vickrey (1955, 1972), Paul (1972), Baumol et al. (1962), and Kahn (1979). One of the major problems involves devising a pricing mechanism that will encourage economic efficiency yet maintain the profitability of the firm.

The issue of efficient pricing for a general multiproduct firm has been examined in the classic article by Baumol and Bradford (1970). They developed a rule of second-best pricing for a firm which would earn negative profits if price were set equal to marginal cost in each market. Their pricing rule maximizes social welfare subject to a constraint which allows the firm to break even. According to the second best rule, the social welfare will be served most effectively not by setting prices equal to or even proportionate to marginal costs but by causing unequal deviations in which the market with elastic demands are priced at levels close to the marginal costs. The prices of the output whose demand is inelastic diverge from the marginal costs by relatively wide margins. A variety of recent work addresses the difficult issue of pricing multiple outputs in budget-constrained firms; see Baumol (1977), Panzar and Willig (1977a), Willig and Baily (1979), and Baumol, Bailey, and Willig (1977).

In what follows, we explicitly rederive the second best rule by using a general formulation in order to clarify its meaning for water supply. It is assumed in this research that the objective of utility commissions or regulatory bodies in the design of rate structure is to maximize community welfare. Equity in income distribution is not an objective of the rate structure in water supply. Therefore, distributional equity associated with the second best pricing problem discussed in Feldstein (1972) is assumed to equal over the outputs of water supply.

Suppose that the water supply firm is able to set the prices of its final outputs but to purchase inputs at prices which remain fixed. The welfare maximizing prices for the water firm facing a budget or profit constraint is written as follows:

Maximize consumer welfare represented by the indirect utility function:

$$V(P_R, P_N, M) \quad (3.79)$$

subject to a budget constraint:

$$P_R Y_R + P_N Y_N - C(Y_R, Y_N; W; Z) = B \quad (3.80)$$

where $V(P_R, P_N, M)$ = indirect utility function³⁵ which is a measure of consumer surplus or welfare,

P_i = price of output i ($i = R, N$)

M = consumer income

$Y_i = Y_i(P_R, P_N, M)$ = demand for output i ($i = R, N$)

$C(Y_R, Y_N; W; Z)$ = total production cost

B = specified surplus or deficit

The Lagrangean for this constrained maximization problem is:

$$L = V(P_R, P_N, M) + \mu [P_R Y_R(P_R, P_N, M) + P_N Y_N(P_R, P_N, M) - C\{Y_R(P_R, P_N, M), Y_N(P_R, P_N, M); W; Z\} - B] \quad (3.81)$$

The basic first-order conditions can be written as:

³⁵ For a complete survey on indirect utility function and duality in consumer theory (and also in production theory), see Diewert (1947a, 1978b).

$$\begin{aligned} \partial L / \partial P_R &= \partial V / \partial P_R + \mu [Y_R + P_R (\partial Y_R / \partial P_R) + P_N (\partial Y_N / \partial P_R) \\ &\quad - (\partial C / \partial Y_R)(\partial Y_R / \partial P_R) - (\partial C / \partial Y_N)(\partial Y_N / \partial P_R)] = 0 \end{aligned} \quad (3.82a)$$

$$\begin{aligned} \partial L / \partial P_N &= \partial V / \partial P_N + \mu [P_N (\partial Y_R / \partial P_N) + Y_N + P_N (\partial Y_N / \partial P_N) \\ &\quad - (\partial C / \partial Y_R)(\partial Y_R / \partial P_N) - (\partial C / \partial Y_N)(\partial Y_N / \partial P_N)] = 0 \end{aligned} \quad (3.82b)$$

$$\partial L / \mu = P_R Y_R + P_N Y_N - C(Y_R, Y_N; W; Z) - B = 0 \quad (3.82c)$$

These expressions (3.82) can be simplified by denoting the price elasticity of demand for outputs as:

$$E_{ip} = (\partial Y_i / \partial P_p) (P_p / Y_i) \quad (3.83)$$

and the marginal cost as:

$$MC_i = \partial C / \partial Y_i \quad (3.84)$$

and further by making use of the result that:

$$\partial V(P_R, P_N, M) / \partial P_i = - Y_i (\partial V(P_R, P_N, M) / \partial M) \quad (3.86a)$$

$$E_{ip} = E_{pi} (P_p Y_p / P_i Y_i) \quad (i, p = R, N) \quad (3.86b)$$

³⁶ it is derived from Roy's identity:

$$Y_i(P_R, P_N, M) = - \frac{\partial V / \partial P_i}{\partial V / \partial M} \quad (3.85)$$

Roy's identity says that the increase in income required by the consumer to compensate it (keep the level V constant) for a price increase is just the extra amount needed to purchase its previously optimal commodity bundle at the new price. For more on Roy's identity and duality theorem in consumer theory, see Diewert (1974a, 1978b).

After rearranging terms, the first-order conditions are then:

$$\mu[1 + E_{RR}(P_R - MC_R)/P_R - E_{RN}(P_N - MC_N)/P_N] = \partial V/\partial M \quad (3.87a)$$

$$\mu[1 + E_{NR}(P_R - MC_R)/P_R + E_{NN}(P_N - MC_N)/P_N] = \partial V/\partial M \quad (3.87b)$$

Setting (3.87a) equal to (3.87b) and rearranging yields a convenient expression for the price-marginal cost markup ratio:

$$\frac{(P_R - MC_R)/P_R}{(P_N - MC_N)/P_N} = \frac{E_{NN} - E_{RN}}{E_{RR} - E_{NR}} \quad (3.88)$$

E_{RR} and E_{NN} are the own price elasticities of demand for residential and nonresidential water, respectively. E_{RN} is the cross price elasticity of demand for residential water (Y_R) with respect to output price of nonresidential water (P_N). When the cross demand elasticities E_{RN} and E_{NR} are zero, (3.88) reduces to the well known "inverse elasticity rule" that in each market, the amount by which price deviates from marginal cost is inversely related to the price elasticity of demand.

Prices which maximize welfare subject to the budget constraint thus must satisfy the following two equations:

$$\frac{(P_R - MC_R)/P_R}{(P_N - MC_N)/P_N} = \frac{E_{NN} - E_{RN}}{E_{RR} - E_{NR}} \quad (3.89a)$$

$$P_R Y_R + P_N Y_N - C(Y_R, Y_N; W; Z) = B \quad (3.89b)$$

Prices which satisfy this budget constrained welfare maximization problem are often referred to as "Ramsey optimal prices," following the

early work of Ramsey (1927) on taxation. Second best or Ramsey optimal prices depend upon marginal costs and demand elasticities, which may require cross-subsidization among consumers. It will be interesting to compare actual prices charged by the water firm with Ramsey prices to determine the degree of cross-subsidization.

It is useful to define budgeted revenue, BR, in order to facilitate interpretation. Budgeted revenue, $BR = C + B$, is simply total cost plus a budgeted surplus or total cost minus a budgeted deficit. Thus, BR is the revenue amount needed to produce the desired surplus or deficit.

Consider a marginal cost pricing system. Unless

$$P_R Y_R + P_N Y_N - BR = 0 \quad (3.90)$$

the budget constraint in (3.89b) will be violated. The marginal cost pricing in this model is only optimal if there are constant returns to scale and no surplus or deficit is budgeted. In most cases, marginal cost pricing does not maximize welfare and some adjustments are needed to satisfy (3.89).

In this model, the budget constraint imposes a certain welfare loss. "Ramsey" optimality requires that prices be set to spread the welfare loss across each output in a way that results in the least welfare loss for all customers. The more unresponsive the demand for an output changes in its price, the less welfare is lost if that price is increased, and so the larger the optimal deviation between price and marginal cost. Thus, the overhead expenses for services produced with economies of scale are best covered with revenues from the more inelastic part of the service.

APPENDIX: DERIVATION OF PARAMETER RESTRICTIONS FOR HOMOGENEITY IN
INPUT PRICES AND OUTPUTS

Homogeneity of degree one in W for fixed Y and Z requires:

$$C(Y, tW, Z) = tC(Y, W, Z) \quad (A.1)$$

Thus,

$$\begin{aligned} \ln C(Y, tW, Z) &= \ln (tC(Y, W, Z)) \\ &= \ln t + \ln C(Y, W, Z) \end{aligned} \quad (A.2)$$

By direct calculation, we therefore obtain:

$$\begin{aligned} \ln C(Y, tW, Z) &= a_0 + \sum_i a_i \ln Y_i + \sum_j b_j (\ln t + \ln W_j) \\ &+ \sum_k c_k \ln Z_k + 1/2 \sum_i \sum_p a_{ip} \ln Y_i \ln Y_p + 1/2 \sum_j \sum_q b_{jq} (\ln t \\ &+ \ln W_j)(\ln t + \ln W_q) + 1/2 \sum_k \sum_r c_{kr} \ln Z_k \ln Z_r \\ &+ \sum_i \sum_j d_{ij} (\ln Y_i (\ln t + \ln W_j)) + \sum_i \sum_k e_{ik} \ln Y_i \ln Z_k \\ &+ \sum_j \sum_k f_{jk} (\ln t + \ln W_j) \ln Z_k \end{aligned} \quad (A.3)$$

Thus

$$\begin{aligned} \ln C(Y, tW, Z) &= \ln C(Y, W, Z) + \sum_j b_j \ln t \\ &+ 1/2 \sum_j \sum_q b_{jq} [(\ln t)^2 + \ln t (\ln W_j) + \ln t \ln W_q] \\ &+ \sum_i \sum_j d_{ij} \ln Y_i \ln t + \sum_j \sum_k f_{jk} \ln t \ln Z_k \end{aligned} \quad (A.4)$$

Now equating (A.4) to (A.2), we can derive relevant parameter restrictions for homogeneous of degree one in input prices. Examining the

second term on the RHS of (A.4), we have

$$\sum_j b_j = 1 \quad (\text{A.5})$$

The third term on the RHS of (A.4) can be written as (after multiplying by 2):

$$(\ln t)^2 \sum_j \sum_q b_{jq} + \ln t \sum_j \sum_q b_{jq} \ln W_j + \ln t \sum_j \sum_q b_{jq} \ln W_q \quad (\text{A.6})$$

For the last two terms of this expression to equal to zero, we must have:

$$\begin{aligned} \sum_q b_{jq} &= 0 \\ \sum_j b_{jq} &= 0 \end{aligned} \quad (\text{A.7})$$

But since b is a symmetric matrix, this implies $\sum_j b_{jq} = \sum_q b_{jq}$ and that

$$\sum_j \sum_q b_{jq} = 0 \quad (\text{A.8})$$

For the translog cost function to be homogeneous of degree one in W , the fourth term on the RHS of (A.4) must be zero. Thus,

$$\ln t \sum_i \sum_j d_{ij} \ln Y_i = 0 \quad (\text{A.9})$$

This requires:

$$\sum_j d_{ij} = 0 \quad (\text{A.11})$$

Similarly, we have restrictions on the last term of (A.4):

$$\sum_j f_{jk} = 0 \quad (\text{A.12})$$

To recapitulate briefly, homogeneity of degree one in input prices W requires the following restrictions:

$$\begin{aligned}
 \sum_j b_j &= 1 \\
 \sum_j b_{jq} &= \sum_q b_{jq} = \sum_j \sum_q b_{jq} = 0 \\
 \sum_j d_{ij} &= 0 \\
 \sum_j f_{jk} &= 0
 \end{aligned}
 \tag{A.13}$$

Since $C(Y, W, Z)$ is symmetric in W and Y , the above argument can be extended to derive the necessary restrictions for homogeneity in outputs Y of an arbitrary degree. Thus,

$$\begin{aligned}
 \sum_i a_i &= \rho \\
 \sum_i a_{ip} &= \sum_p a_{ip} = \sum_i \sum_p a_{ip} = 0 \\
 \sum_i d_{ij} &= 0 \\
 \sum_i e_{ik} &= 0
 \end{aligned}
 \tag{A.14}$$

CHAPTER 4

EMPIRICAL METHODOLOGY

The purpose of this chapter is to provide an econometric framework necessary for empirical implementation of the model developed in Chapter 3. Section 4.1 is intended to give a statistical formulation of the translog cost model. This section points out some of the problems of using the ordinary least squares method of estimating the parameters of the translog cost function. Then, the section demonstrates the use of the Iterative Zellner Efficient (IZEF) method which jointly estimates the translog cost function and the input share functions. The IZEF method is asymptotically equivalent to the maximum likelihood estimation method. Section 4.2 discusses the data used in the analysis including problems resulting from the lack of good data concerning water utilities and the source of the data. Finally, Section 4.3 presents some sample statistics of the data such as means and standard deviations.

4.1 ECONOMETRIC SPECIFICATION AND ESTIMATION PROCEDURE

The translog cost function as shown in equation (3.10) in Chapter 3 is not in a form so that it can be estimated econometrically. Therefore, this section provides an econometric formulation of the translog cost model and presents a method of estimating the parameters. We rewrite the translog cost function (3.10) as:

$$\begin{aligned}
 \ln C = & a_o + a_R \ln Y_R + a_N \ln Y_N + b_L \ln W_L + b_K \ln W_K \\
 & + b_E \ln W_E + c_U \ln Z_U + c_D \ln Z_D + 1/2 a_{RR} (\ln Y_R)^2 \\
 & + a_{RN} \ln Y_R \ln Y_N + 1/2 a_{NN} (\ln Y_N)^2 + 1/2 b_{LL} (\ln W_L)^2 \\
 & + b_{LK} \ln W_L \ln W_K + b_{LE} \ln W_L \ln W_E + 1/2 b_{KK} (\ln W_K)^2 \\
 & + b_{KE} \ln W_K \ln W_E + 1/2 b_{EE} (\ln W_E)^2 + 1/2 c_{UU} (\ln Z_U)^2 \\
 & + c_{UD} \ln Z_U \ln Z_D + 1/2 c_{DD} (\ln Z_D)^2 + d_{RL} \ln Y_R \ln W_L \\
 & + d_{RK} \ln Y_R \ln W_K + d_{RE} \ln Y_R \ln W_E + d_{NL} \ln Y_N \ln W_L \\
 & + e_{NK} \ln Y_N \ln W_K + d_{NE} \ln Y_N \ln W_E + e_{RU} \ln Y_R \ln Z_U \\
 & + e_{RD} \ln Y_R \ln Z_D + e_{NU} \ln Y_N \ln Z_U + e_{ND} \ln Y_N \ln Z_D \\
 & + f_{LU} \ln W_L \ln Z_U + f_{LD} \ln W_L \ln Z_D + f_{KU} \ln W_K \ln Z_U \\
 & + f_{KD} \ln W_K \ln Z_D + f_{EU} \ln W_E \ln Z_U + f_{ED} \ln W_E \ln Z_D \\
 & + E_C
 \end{aligned} \tag{4.1}$$

The disturbance term E_C is added in (4.1) to account for the fact that the translog function is an approximation of the underlying true function.

The translog cost function (4.1) entails no a priori restrictions on production characteristics. Scale economies can not only be increasing, but can vary for different levels of output. The elasticity of substitution varies between inputs and for different levels of input usage. The same is true for the elasticity of transformation among outputs.

It is necessary to estimate the unknown parameters of the translog cost function. The difficulty, however, is that there are a large number of regressors even for a relatively small number of outputs and inputs. If there are any substantial degree of multicollinearity among regressors, estimates of the parameters by the single equation (ordinary least squares and maximum likelihood) estimators of (4.1), although attractive from the point of view of simplicity, may be imprecise. This is particularly true unless there are significant variations in outputs, input prices, and operating variables across the sample. The multicollinearity problem is explicitly taken into account in the estimation of the translog function in the following manners.¹

First, the regressors are normalized to have a mean of one and a standard deviation of zero, as indicated in the last chapter. It has been suggested (Marquardt and Snee 1975, Marquardt 1980) that normali-

¹ The imposition of cost minimization conditions (such as linear homogeneity of input prices as shown below) which place restrictions across parameters helps circumvent the multicollinearity which may be present.

zation or standardization reduces the high correlations among variables for the model involving quadratic or higher order terms. Normalization of variables also facilitates interpretation of the coefficients by making the effects of the regressor variables, over their actual ranges in the data, of comparable magnitude.

Second, we use additional information contained in the cost share equations in order to increase the efficiency of the parameter estimates.² Then we can jointly estimate the cost function and cost share equation as a system of equations.³

Differentiating the translog cost function with respect to $\ln W_L$, $\ln W_K$, and $\ln W_E$, the cost share equations take the following form:

$$S_L = b_L + b_{LL} \ln W_L + b_{LK} \ln W_K + b_{LE} \ln W_E \\ + d_{RL} \ln Y_R + d_{NL} \ln Y_N + f_{LU} \ln Z_U + f_{LD} \ln Z_D + E_L \quad (4.2a)$$

$$S_K = b_K + b_{LK} \ln W_L + b_{KK} \ln W_K + b_{KE} \ln W_E \\ + d_{RK} \ln Y_R + d_{NK} \ln Y_N + a_{KU} \ln Z_U + f_{KD} \ln Z_D + E_K \quad (4.2b)$$

² If the assumption of marginal cost pricing is tenable, it is possible to use the information given by revenue share equations in addition to the cost share equations, as have been done by Hall (1973), Burgess (1974), Darrough and Heneike (1978), and Denny, Fuss, Everson, and Waverman (1981).

³ Guilkey and Lovell (1980), however, provide a contrary argument. Using the monte carlo experiments, they found the single equation estimator to perform marginally better than the system estimator when the underlying technology is not translog. Note the qualification when the underlying technology is not translog. We assume, however, that the underlying technology is unknown to us.

$$S_E = b_E + b_{LE} \ln W_L + b_{KE} W_K + b_{EE} \ln W_E \\ + d_{RE} \ln Y_R + d_{NE} \ln Y_N + f_{EU} \ln Z_U + f_{ED} \ln Z_D + E_E \quad (4.2c)$$

The disturbance terms E_j ($j = L, K, E$) are placed at the end of each S_j ($j = L, K, E$) to account for random errors in optimization behavior in minimizing costs.

In this manner, we obtain three cost share equations in addition to the cost function, but have not introduced any additional unknown parameters. We assume that the disturbances in the cost function (4.1) and the cost share equations (4.2) have joint normal distribution⁴ but allow for contemporaneous correlation across equations since they affect all cost share functions simultaneously. This indicates that we allow nonzero correlation for a particular firm but impose zero correlation across firms.

Since a well behaved cost function is linear homogeneous in input prices, the cost shares sum to unity, i.e., $S_L + S_K + S_E = 1$, for all input prices, outputs, and other variables. Thus, we can obtain the following restrictions on the parameters of the cost function and share equations:

⁴ Woodland (1979) noted the inappropriateness of the normality assumptions with the system of share equations. According to him, the normality is inconsistent with the properties of the cost shares S_j . The normality assumption requires that the random variable be continuously distributed from $-\infty$ to $+\infty$, where there is a positive probability that the shares will be outside the 0-1 range. This is inconsistent with the requirement that each share is discretely distributed between zero and one. The normal distribution also assumes that the shares are symmetrically distributed. This seems unlikely for shares with expected values near zero or one. However, Woodland also demonstrated that correcting the normality assumption makes little difference in the estimation results.

$$\begin{aligned}
b_L + b_K + b_E &= 1 \\
b_{LL} + b_{LK} + b_{LE} &= 0 \\
b_{LK} + b_{KK} + b_{KE} &= 0 \\
b_{LE} + b_{KE} + b_{EE} &= 0 \\
d_{RL} + d_{RK} + d_{RE} &= 0 \\
d_{NL} + d_{NK} + d_{NE} &= 0 \\
f_{LU} + f_{KU} + f_{EU} &= 0 \\
f_{LD} + f_{KD} + f_{ED} &= 0
\end{aligned} \tag{4.3}$$

The fact that the cost share equations sum to unity poses the estimation problem present in equations (4.1) through (4.2). The problem arises because the parameters of any two share equations together with the parameters of the cost function exactly identify all parameters of the system. This implies that the sum of the disturbances across the share equations is zero for each firm. This further implies that the share disturbance variance-covariance matrix is singular and nondiagonal. Estimation techniques for a system of equations with singular variance-covariance matrix and with constraints across equations due to the symmetry restrictions are now well known (see Theil 1971 Section 6.7). In the estimation, we delete one cost share equation. We then jointly estimate the remaining two cost share equations together with the translog cost function.

Deleting arbitrarily the capital share equation, we can rewrite the above restrictions in the following way:

$$b_K = 1 - b_L - b_E$$

$$b_{LK} = -b_{LL} - b_{LE}$$

$$b_{KK} = b_{LL} + b_{EE} + 2b_{LE} \quad (4.4)$$

$$b_{KE} = b_{LE} - b_{EE}$$

$$d_{EK} = -d_{RL} - d_{RE}$$

$$d_{NK} = -d_{NL} - d_{NE}$$

$$f_{KU} = -f_{LU} - f_{EU}$$

$$f_{KD} = -f_{LD} - f_{ED}$$

Now substituting the restrictions into the cost function and the share equations, we can obtain the system of equations shown below. The translog cost function to be estimated becomes:

$$\ln C - \ln W_K = a_0 + a_R \ln Y_R + a_N \ln Y_N + b_L (\ln W_L - \ln W_K)$$

$$+ b_E (\ln W_E - \ln W_K) + c_U \ln Z_U + c_D \ln Z_D$$

$$+ 1/2 a_{RR} (\ln Y_R)^2 + a_{RN} \ln Y_R \ln Y_N + 1/2 a_{NN} (\ln Y_N)^2$$

$$+ b_{LL} [1/2 (\ln W_L)^2 + 1/2 (\ln W_K)^2 - \ln W_L \ln W_K]$$

$$+ b_{EE} [1/2 (\ln W_E)^2 + 1/2 (\ln W_K)^2 - \ln W_K \ln W_E]$$

$$+ b_{LE} [(\ln W_K)^2 + \ln W_L \ln W_E - \ln W_L \ln W_K \ln W_E]$$

$$\begin{aligned}
& + 1/2c_{UU}(\ln Z_U)^2 + c_{UD} \ln Z_U \ln Z_D + 1/2c_{DD}(\ln Z_D)^2 \\
& + d_{RL}(\ln W_L - \ln W_K) \ln Y_R + d_{RE}(\ln W_E - \ln W_K) \ln Y_R \\
& + d_{NL}(\ln W_L - \ln W_K) \ln Y_N + d_{NE}(\ln W_E - \ln W_K) \ln Y_N \\
& + e_{RU} \ln Y_R \ln Z_U + e_{RD} \ln Y_R \ln Z_D \\
& + e_{NU} \ln Y_N \ln Z_U + e_{ND} \ln Y_N \ln Z_D \\
& + f_{LU}(\ln W_L - \ln W_K) \ln Z_U + f_{EU}(\ln W_E - \ln W_K) \ln Z_U \\
& + f_{LD}(\ln W_L - \ln W_K) \ln Z_D + f_{ED}(\ln W_E - \ln W_K) \ln Z_D \\
& + E_C
\end{aligned} \tag{4.5}$$

The cost share equations to be estimated become:

$$\begin{aligned}
S_L = & b_L + b_{LL}(\ln W_L - \ln W_K) + b_{LE}(\ln W_E - \ln W_K) \\
& + d_{RL} \ln Y_R + d_{NL} \ln Y_N + f_{LU} \ln Z_U + f_{LD} \ln Z_D + E_L
\end{aligned} \tag{4.6a}$$

$$\begin{aligned}
S_E = & b_E + b_{LE}(\ln W_L - \ln W_K) + b_{EE}(\ln W_E - \ln W_K) \\
& + d_{RE} \ln Y_R + d_{NE} \ln Y_N + f_{EU} \ln Z_U + f_{ED} \ln Z_D + E_E
\end{aligned} \tag{4.6b}$$

Equations (4.5), (4.6a), and (4.6b) comprise the constrained multivariate regression system to be estimated. Since the disturbances are correlated across equations, the two stage estimation method suggested by Zellner (1962, 1963) yields asymptotically more efficient estimates than the ordinary least squares method. However, the estimates obtained by Zellner's two-stage least squares to the con-

strained multivariate regression may be sensitive to the equation deleted. Therefore, the Iterative Zellner Efficient (IZEF) method⁵ can be utilized, in which the estimates are iterated until the estimated weighed disturbance variance-covariance matrix converges to an identity matrix.

The IZEF method is invariant to the cost share equation deleted and asymptotically equivalent to the maximum likelihood estimation method; see Barten (1969), Kmenta and Gilbert (1968), and Oberhofer and Kmenta (1974). The iterative procedure of the Zellner method is important since it is the assumption of normality, meaning in part covariance values of zero, that makes the IZEF technique the same as the maximum likelihood estimator. That is, the process of iteration forces the covariance close to zero (assuming convergence), leading to the claim that the IZEF technique is equivalent to the maximum likelihood approach. The Zellner efficient technique is useful for this analysis only if it is iterative. A computation algorithm widely used is that of Berndt et al. (1974). This algorithm essentially uses the Gauss-Newton method of nonlinear least squares (see Bard 1974).

Since we obtain maximum likelihood estimates, we are able to test hypotheses such as homogeneity and separability by using the likelihood ratio test. The likelihood ratio is the ratio of the maximum value of the likelihood function L for the model subject to restrictions (ω) to the maximum value of the likelihood function for the model without re-

⁵ It is sometimes referred to as Malinvaus's (1981 Chapter 9) iterative minimum distance estimation method.

strictions (ω). Hence we can write the likelihood ratio as:

$$\lambda = \max \mathcal{L}(\omega) / \max \mathcal{L}(\Omega) \quad (4.7)$$

We test hypotheses using the fact that the test statistic⁶ $-2 \ln \lambda$ has an asymptotic χ^2 distribution with degrees of freedom equal to the number of independent restrictions being imposed. The null hypothesis is rejected if calculated $-2 \ln \lambda$ exceeds χ^2 for the chosen value of the significance level. The likelihood ratio test is a more stringent test of significance than the commonly used t-test and F-test since the covariance matrix is not held fixed for the restricted estimation (Berndt and Savin 1977).

Some caveats are in order with regard to the direct use of the IZEF method. For a sample consisting of time series cross section data or of the short period of individual observations, it may not be unreasonable to expect autocorrelated disturbances. Berndt and Savin (1975) and Beach and MacKinnon (1979) have shown that the singularity of the contemporaneous disturbance covariance matrix of the full cost share and cost equation system imposes restrictions on the autoregressive process. If the model estimated does not satisfy these restrictions, then maximum likelihood estimation will not be invariant to the deleted cost share equation. A consequence of this, of course, is that the likelihood ratio

⁶ For normally distributed disturbances, this can be written as:

$$-2 \ln \lambda = T(\ln |\hat{\Sigma}_{\omega}| - \ln |\hat{\Sigma}_{\Omega}|) \quad (4.8)$$

where $|\hat{\Sigma}_{\omega}|$ and $|\hat{\Sigma}_{\Omega}|$ are the determinants of the estimated error covariance matrices for the restricted and unrestricted models, and T is the number of observations.

test statistic will also be affected by the cost share equation deleted.

A simple specification for this problem is to assume a first order stochastic process:

$$E_t = \rho E_{t-1} + u_t \quad (4.9)$$

where E_t is an (3×1) vector of random disturbances consisting of E_C and E_j ($j = L, E$), and u_t is an (3×1) vector of independently identically distributed normal random vectors. The ρ is the same for each equation in order for the estimates to be invariant with respect to the choice of omitted equation.

The IZEF method is asymptotically equivalent to the maximum likelihood method under the assumption of no heteroscedasticity within equations. When estimating the equations of the multivariate regression system, we can account for within-equation heteroscedasticity using the following procedure. First, each equation in the system is estimated using ordinary least squares. The resulting regression residuals, which we can label ϵ_i^2 , can then be used to obtain consistent estimates of the error variances of the i th equation σ_i^2 :

$$\hat{\sigma}_i^2 = \frac{1}{T - k_i - 1} \sum_{t=i}^T \epsilon_{it}^2 \quad (4.10)$$

where T is the number of observations and k_i is the number of independent variables in the i th equation. Different estimates of $\hat{\sigma}_i^2$ will be of course obtained for each equation in the system. The data can then be transformed by dividing each observation by an appropriate estimated error term standard deviation $\hat{\sigma}_i$, and the entire system of equations

is re-estimated using the IZEF estimation.

The translog parameter estimates will be consistent under the assumption that all explanatory variables in the cost function (4.1) and the cost share equations (4.2) are exogeneous. However, when dealing with aggregate data, it may be inappropriate to assume that input price and quantities are exogeneous⁷ and that regressors in both equations are uncorrelated with the disturbances (see Berndt and Wood 1975). This introduces a problem of simultaneous equation bias in the parameter estimates. To provide consistent estimates of the parameter estimates, an efficient procedure would use the method of instrumental variables by using the Iterative Three-Stage Least Squares (I3SLS) estimator. The I3SLS estimator is asymptotically equivalent to the Full-Information Maximum Likelihood (FIML) estimator.

A problem with the I3SLS estimation is that there are no obvious choices for instrumental variables to remove the simultaneity problem. There is some evidence, however, though not unanimous, that the estimation results are insensitive to the problem. For instance, Berndt and Christensen (1973) and Appelbaum (1979a) found little differences in the results obtained from IZEF and I3SLS techniques.⁸ The former is not

⁷ If the input prices are endogenous, or $W = W(Y)$, the cost function is $C(Y, W, Z) = C(Y, W(Y), Z)$. Then even under constant returns to scale, $C(Y, W, Z) = Y \times C(W, Z)$ does not hold as is commonly assumed. The consequence of assuming that input prices are not exogenous and therefore not constant to the firm is that an underlying supply curve for each input must be specified. In practice, this is not usually done.

⁸ However, there is a possibility that this may have occurred because the instruments chosen failed to purge the endogeneity of the input prices rather than because endogeneity was unimportant, as they suggested.

designed to take account of simultaneity problems, while the latter is.

4.2 SOURCE AND CONSTRUCTION OF DATA

A major obstacle to meaningful studies on water utilities is the presence of good data. One source of data is the AWWA Statistical Report published by the American Water Works Association. The Report is published about every five years and contains operating data from a large sample of utilities. Another source is the U.S. Environmental Protection Agency (EPA) data base. The EPA recently conducted a series of studies on selected water utilities to determine the economics of water supply. Initially, the data were collected for ten years for each utility in EPA's ten regions for major operating and capital cost categories.

Both AWWA and EPA data are far from adequate for this study of U.S. water utilities. Although the AWWA data seem abundant, it only contains sufficient information for studies on water demand, while the present study is clearly geared to supply issues. The problem with the EPA data is the lack of complete records; some utilities have information on some variables, some do not, and consistent time series data are not available.

Regretably, reasonable alternatives do not exist. We therefore chose the EPA data as the basis of our study on water utilities. The time period chosen is the calendar year 1973, because of the availability of relevant data for that year. The sixty utilities chosen represent a technical and geographical cross-section at the United States. Most utilities are municipally owned, providing their services to a metro-

politan area.

Since some critical data are not readily available for each utility, estimations had to be made. The estimation of the missing data is far from perfect, but it is impossible to determine the extent to which the estimates bias our results until better data become available.

Total input cost and input shares are readily computed. Total cost (C) is computed for each utility by total labor costs plus total capital costs plus total energy costs. The cost shares (S_j) are calculated by dividing the cost attributable to each input by total cost.

Labor costs are the amount of gross payroll for each utility. Energy costs are total expenditures on power paid by each utility. Capital costs are assumed as interests and depreciation allowances on the plant in service. Depreciation is the loss in value of property due to physical deterioration of the plant and sometimes due to some factors external to the operation of the water utility plant. Depreciation allowances are based on annual percentage of the items of plant investment. The annual percentages, or rates of depreciation, are established at levels considered necessary to yield a dollar amount equivalent to the initial plant investment by the end of the usual life of the physical property. Water utilities generally use depreciation funds as a source of capital for replacements. Interest is the dollar amount paid in a given year on any borrowing mechanisms for a utility. It may be noted that depreciation is based on the historical cost of the facility and not on the costs required to reproduce the facility. Therefore, capital costs as constructed may not represent the replacement costs which are the more relevant concept, and thus can be

considered as approximate, not true, costs of capital.

Water is measured in terms of the amount of water treated. Raw or treated water is a more representative measure of water output than revenue-producing water in order to gain insight into the scarcity of water. The difference between the two is the amount of water loss during transmission to consumers. Water is broken into residential (Y_R) and nonresidential (Y_N) uses. Residential water is the water delivered to residencies for the purpose of normal living and includes all single and multi-family dwelling units and apartments. Nonresidential water is the water delivered to industrial, commercial, wholesale, and other users. It includes schools, churches, fire stations, and other public institutions. Residential customers constitute the majority of public water users. Not all utilities in the sample, however, have a detailed breakdown of the types of customers they serve, so this categorization is often gross. Yet, dividing water usage into residential and nonresidential consumption will provide some basis for analyzing demand and cost patterns among different types of customer classes.

The quantity of labor is measured in yearly man hours (assumed to be equal to the number of working hours of persons employed in each utility to produce water during the given year). The price of labor per man-hour (W_L) may be obtained by dividing the gross payroll by the number of yearly man-hours. If the price of labor was not available in a given utility, the average in the EPA region in which the utility is principally located was used. The construction of the price of labor assumes a high degree of homogeneity in labor. But we have labor of

varying skills and educations. This introduces an aggregation problem,⁹ which is not a main concern in this research.

One of the most difficult pieces of operating data to obtain is energy use and energy price data. Although energy costs are available in most utilities, few utilities have considered them enough, until now, to document separately and use as a management tool for evaluating the efficiency of utility operations.

The price of energy (W_E) can be estimated by dividing total energy expenditures by yearly KWH usage. However, the estimation was troublesome, since most information on the price of energy was missing. The Department of Energy publishes a statistical book of electric utilities in the United States which includes operating data. The revenue from the sale of electricity was divided by the total number of Kilowatt-hours provided by the electric utility in the region in which the water utility is principally located. This ratio was used as an average price of energy for each water utility. This approach provides a correct estimation only in the case where all electricity is priced uniformly regardless of the user or the amount purchased. With the presence of declining block rates practiced by electric utilities, however, the constructed energy price can be considered as an approximate (and biased) measure of the true energy price.

Determining the price of capital (W_K) is considerably less straightforward than determining the price of labor or energy. The price

⁹ For aggregation problem and index numbers, see Samuelson and Swamy (1974), Diewert (1976, 1978), Lau (1979), Theil (1973), Hulten (1973), etc.

of capital is formulated in terms of rental price of capital as suggested by Jorgenson (1963) and Christensen and Jorgenson (1969, 1970). Under this formula, the price of capital is computed by considering prevailing interest rates as an opportunity cost for investment, then accounting for the incidence of taxes, depreciation, and inflation. Empirical studies that employ versions of this formula include Berndt and Christensen (1973) and Berndt and Wood (1975).

Since we are dealing with municipal utilities for cross sectional samples, we can disregard taxes and inflation. The imputed rental price of capital for a given utility can be estimated by:

$$W_K = (r + d)NER \quad (4.11)$$

where r is the bond rate on long-term debts, d is the depreciation rate, and NER is the Engineering News-Record construction costs. Under regulatory commission control, the depreciation rate is fixed or approved by order and is adjusted from time to time to accomplish their objective. The price of capital as constructed is far from ideal and may not reflect the true current dollar worth or the contribution of capital to the production process.

The service distance (Z_D) is the total miles of transmission and distribution pipe (as recorded from annual reports, in the water delivery system. Capacity utilization (Z_U) represents the relationship between the average rate of plant usage and maximum (designed or rated) capacity. Capacity utilization rate in this research is measured by the load factor for a water system. The load factor measures the average use of system

capacity as a percent of maximum demand, i.e., the ratio of average demand (the sum of residential and nonresidential usage of water) to peak demand. The load factor is defined with reference to a specific time period, i.e., month. The relationship of the capacity utilization rate to the load factor for a specific water system is determined by the extent of reserve capacity.

4.3 ANALYSIS OF THE SAMPLE DATA

The sample data set chosen for this study consists of 60 utilities which are well distributed geographically throughout the United States. The sample shows wide variations in the major variables which reflect differences in the operation of water utilities. These utilities range from a very large system treating 373 million gallons of water per day to a very small system treating 16 thousand gallons of water per day.

Sample means and standard deviations for the variables used in this research are tabulated in Table 4.1. Total cost is expressed in terms of thousands of dollars. The price of labor is expressed in terms of dollars per hour, and the price of energy is in terms of dollars per kilowatt per hour (KWH). The capital price is an index compounded of interest rate, depreciation rate, and construction costs.

The majority of water is delivered to residential users (Y_R), constituting 78% of total water treated. 22% of the water is delivered to nonresidential customers (Y_N) such as industrial, commercial, and wholesale users as well as schools, churches, fire stations, etc. In general, costs of supplying water to residential users are higher than the cost of supplying water to nonresidential users, mainly because of

Table 4.1
Sample Statistics

<u>Variables</u>	<u>Notation</u>	<u>Mean</u>	<u>Std Dev</u>
Total Cost (thousands \$)	C	2379.456	4632.01
Labor Share	S_L	0.418	0.16
Capital Share	S_K	0.465	0.17
Energy Share	S_E	0.117	0.11
Residential Water (MGD)	Y_R	20.281	39.28
Nonresidential Water (MGD)	Y_N	10.992	27.74
Labor Price (\$/hr)	W_L	4.068	1.12
Capital Index	W_K	145.157	30.53
Energy Price (¢/KWH)	W_E	0.796	0.47
Capacity Utilization	Z_U	0.563	0.17
Service Distance (Miles)	Z_D	539.485	1003.49

pipe and billing costs. Standard deviations of Y_R and Y_N are higher than the respective means, reflecting the variation in utility size across the sample. This is also shown by the high standard deviation of the total cost.

Labor and capital constitute about equal proportions in the overall utility's budget. The share of energy in total cost is small relative to those of labor and capital. Energy is, however, an increasingly important factor in the operation of water utility. Energy expenditures are directly related to the actual level of production of water and essential for treatment of water and in the distribution system. Operating expenditures on labor and energy for an average utility are as high as capital costs.

Average wage rate in the sample is about 4 dollars per hour with a standard deviation of about one dollar. The variation in the wage rate reflects differences in salaries of operators, managers, and secretaries. Capital price constructed as a composite index shows some variation among the chosen sample. Mean energy price in the sample is about \$0.80 per KWH.

Capacity utilization as measured by the load factor for water systems is on the average about 60%, thus 40% of the capacity remains idle in the operation of the average water utility. The average mile of pipe in system for all utilities is 539 miles, but there is a wide variation in the service distance across utilities as shown by the high standard deviation.

CHAPTER 5

EMPIRICAL FINDINGS

This chapter presents the empirical results generated via the previously detailed theoretical methodology. Section 5.1 summarizes the parameter estimates of the translog cost function. Alternative translog cost models are estimated to see if the estimated cost function is robust with respect to other restrictive specifications of the technology of water supply. In Section 5.2, elasticities of substitution among labor, capital, and energy, and corresponding own and cross price elasticities of input demand are analyzed. The input demand with respect to outputs and operating variables is also discussed. Section 5.3 investigates the question of scale economies in water supply. Section 5.4 discusses the impact of changes in operating variables, capacity utilization and service distance, on costs. Section 5.5 presents estimates of marginal costs and then analyzes the transformation function of water supply. Finally, Section 5.6 concludes the chapter with an examination of pricing rules for residential and nonresidential water supply and the issue of cross-subsidization among user classes.

5.1 ESTIMATED COST FUNCTIONS AND THEIR PROPERTIES

As noted earlier, the methodology of this research is based on the translog multiproduct cost specification which imposes a few a priori restrictions on the structure of production. Further, these restrictions are imposed as testable hypotheses rather than as maintained hypotheses. Based on these tests, the best model is chosen.

Parameter estimates and asymptotic standard errors, together with the pseudo- R^2 , for the translog cost function (4.1) are given in Table 5.1. Parameters are estimated by use of the iterative Zellner efficient method which jointly estimates the translog cost function and cost share equations. Parameter values are interpreted as normalized responses of independent variables since independent variables are normalized with means 1 and standard deviations 0.

An appropriate measure of goodness of fit in this multi-variate regression system is the "pseudo- R^2 ", which states the proportion of generalized variance in the system of equations explained by variation in the right-hand variables.¹ The pseudo- R^2 is calculated as $1 - |\Sigma_1|/|\Sigma_2|$, where $|\Sigma_1|$ is the determinant of the estimated residual covariance matrix and $|\Sigma_2|$ is the determinant when all of the slope coefficients are constrained to be zero. This systems-measure of fit collapses to the conventional R^2 when there is but a single equation. A degree of freedom

¹ It may be noted that for estimation procedures other than ordinary least squares, R^2 cannot in general be interpreted as the proportion of variance explained, since the residuals are no longer orthogonal to the regressors. However, it still provides a useful indicator of goodness of fit.

Table 5.1

Parameter Estimates for the General Translog Multiproduct Cost Function
(Asymptotic Standard Errors in Parentheses)

a_o 7.71495 (0.16594)	b_{LK} -0.00392 (0.07483)	d_{NK} 0.01496 (0.03132)
a_R 0.53037 (0.41870)	b_{LE} -0.03843 (0.03838)	d_{NE} 0.00686 (0.01944)
a_N 0.25741 (0.27472)	b_{KK} -0.00409 (0.08049)	e_{RU} -0.59305 (0.46594)
b_L 0.46228 (0.03111)	b_{KE} -0.00018 (0.03809)	e_{RD} 0.44056 (0.34221)
b_K 0.45145 (0.33986)	b_{EE} 0.03860 (0.03502)	e_{NU} 0.13495 (0.23823)
b_E 0.08626 (0.02114)	c_{UU} 0.38587 (0.48304)	e_{ND} -0.32537 (0.25618)
c_U -0.03123 (0.59445)	c_{UD} 0.44037 (0.33294)	f_{LU} -0.01873 (0.04443)
c_D 0.24798 (0.27514)	c_{DD} -0.08081 (0.15927)	f_{LD} -0.04175 (0.03123)
a_{RR} -0.42487 (0.65642)	d_{RL} 0.08885 (0.04256)	f_{KU} -0.00304 (0.04865)
a_{RN} 0.10086 (0.36986)	d_{RK} -0.11238 (0.04661)	f_{KD} 0.09794 (0.03449)
a_{NN} 0.14868 (0.27540)	d_{RE} 0.02353 (0.02870)	f_{EU} 0.02177 (0.02993)
b_{LL} 0.04235 (0.08434)	d_{NL} -0.02182 (0.02887)	f_{ED} -0.05619 (0.02127)
Pseudo- R^2 0.88980		

correction has been applied of $(GT-K)/GT$, where G is the number of equations, T is the number of observations, and K is the number of coefficients estimated. The value of the pseudo- R^2 is invariant to the choice of equation to be omitted from the system. The pseudo- R^2 is high, so a large proportion of the variation in the data is explained by the translog model.

As can be seen from Table 5.1, a large proportion of estimated coefficients on the whole do not achieve statistical significance at the conventional level of significance.² This problem has been encountered in this kind of study using the flexible functional form. One rational explanation must come from the fact that there are a large number of independent variables in the system which burden the estimation of the model and cause this problem to occur. Another explanation must come from the nature of the data being used. This is particularly true for cross section data. Cross section samples often conceal a great deal of variability among firms because of their presumed heterogeneity. This variability among firms will be reflected in the large standard errors of

² Conventional practice in most applications has been to use some form of significance test (usually the 5 percent significance level) in interpreting statistical estimates of parameters. However, in recent work with the flexible functional forms for production and cost functions, this practice has not been entirely uniform. In the case of translog production function estimates, derivation of an estimated covariance structure becomes subject to considerable question. With the translog cost function, the ratios of estimated coefficients to asymptotic standard errors have been used. However, often the significance tests are not used because of the lack of information on the properties of their ratios; see Berndt and Wood (1975a, 1979), Pindyck (1979), Griffin and Gregory (1977). Yet, when necessary, these ratios are interpreted to have asymptotic t-distributions. This is the interpretation used in this study.

the estimated coefficients.

Still another explanation is the quality of data used. A meaningful study on water utilities requires the existence of good data. However, as noted before, the data chosen suffer from some weaknesses and a potential measurement error problem. The measurement error problem occurs because the price indices constructed for energy and capital are by no means ideal and they require improvement. Because of these shortcomings, the data may include observations that are too far from the point of approximation of the translog function to the true cost function. Thus, some discrepancies in the estimated coefficients and large standard errors will result. In the face of limitations imposed by the data currently available, however, the best we can hope is for more attention to be devoted in the future to obtaining higher quality data.

Four alternative translog cost models are estimated: (1) constant returns to scale (CRTS); (2) separable; (3) nonjoint; and (4) Cobb-Douglas. These restricted models impose restrictions on the parameters of the general translog cost function.

Constant returns to scale are obtained by imposing restrictions on the translog cost model such that:

$$a_R + a_N = 1$$

$$a_{NN} = a_{RR}$$

$$e_{NU} = -e_{RU}$$

$$e_{ND} = -e_{RD}$$

The separable model is obtained by imposing restrictions:

$$\begin{aligned}
 d_{RL} &= 0 & d_{NL} &= 0 \\
 d_{RK} &= 0 & d_{NK} &= 0 \\
 d_{RE} &= 0 & d_{NE} &= 0
 \end{aligned}$$

The nonjoint model is obtained by imposing restrictions:

$$a_{RN} + a_R a_N = 0$$

The Cobb-Douglas functional form is a special case of the translog form obtained by:

$$C = a_o Y_R^{a_R} Y_N^{a_N} W_L^{b_L} W_K^{b_K} W_E^{b_E} Z_U^{c_U} Z_D^{c_D}$$

In order to test for different restricted models, the systems of equations are estimated with and without the corresponding restrictions imposed and the likelihood ratio tests can be used. The test statistics for different restricted models are presented in Table 5.2. We found all four restricted models are rejected at the .05 level of significance. Thus, the full generality of the general translog cost model is required to adequately represent the structure of water supply. This has significant implications for the existing methodology of specification of the technology of water supply.

As indicated previously in Chapter 2, the methodologies currently most in use are based on a number of restrictive assumptions on technology such as input-output separability and nonjoint production. Also, previous studies employ the extremely simplified Cobb-Douglas functional form. The Cobb-Douglas form is unnecessarily restrictive,

Table 5.2
Test Statistics for Restricted Models

Models	Number of Restrictions	Test Statistics	Critical $\chi^2(.05)$
CRTS	4	11.35	5.99
Separability	6	15.09	12.59
Nonjointness	1	9.21	3.84
Cobb-Douglas	20	59.22	31.40

since second order terms which give information on substitution possibilities among inputs are missing. The Cobb-Douglas form is rejected in this test as an appropriate functional representation for the cost function of water supply.

Since the input-output separability condition is rejected, this implies that it is preferable to disaggregate outputs into several components instead of employing a single output measure. Disaggregation of outputs would provide greater depth to the analysis of water supply. Since nonjointness is rejected, estimating the cost function separately for residential and nonresidential water supply would give erroneous information. It is more appropriate to estimate the cost function jointly for residential and nonresidential users. In sum, the translog cost model utilizing the multiproduct production process is a complete description of the technology of water supply. No arbitrary restrictions on the technology are required.

Furthermore, in order to test whether the inclusion of operating variables improves the specification of the model in a statistically meaningful fashion, all the coefficients associated with the capacity utilization and service distance variables were set to zero and the likelihood ratio test was applied. The hypothesis that the operating variables should not be used in the model was rejected at the .05 level of significance. This indicates that there exist significant non-homotheticities in the water supply technology with respect to outputs, input prices, and operating variables. Thus, the inability to account for changes in capacity utilization and service distance would lead to

spurious inferences about the structure and cost of water supply.³

The estimated cost function must satisfy some regularity conditions. However, since the translog cost function does not satisfy the regularity conditions globally, it is necessary to check local properties at each data point. Linear homogeneity in input prices was originally imposed in the estimation. The monotonicity is satisfied since the fitted input shares are all positive at each data point. The local convexity of the estimated translog cost function was checked by computing eigenvalues of the Hessian matrix. The eigenvalues for the Hessian matrix evaluated at sample means are all negative, thus satisfying negative definiteness. These conditions are also satisfied in most sample observations. We can therefore conclude that the estimated translog cost function is well-behaved.

5.2 STRUCTURE OF INPUT DEMAND

A concise description of technology of the firm is provided by the elasticities of substitution and the elasticities of input demand. In this section, estimates of input responses as input prices change are presented. Then, the input requirements as demand for water (residential and nonresidential) and operating variables change are discussed. There have not been any attempts in the past at estimating the input demand functions for water supply.

³ Earlier in Kim (1980), a simple translog cost model was estimated in which output was not disaggregated into customer classes and operating variables were not included. This model did not explain the behavior of utilities in view of the methodology of the present research.

Elasticities of Substitution and Input Price Elasticities

The elasticities of substitution among inputs describe the effect of input prices on factor intensities. The elasticities of substitution are also a measure of the essentiality of economic resources. A resource is said to be essential if production cannot continue in its absence. These elasticities are nonlinear functions of the estimated parameters, since the cost shares are themselves functions of the parameters. The best summary evidence concerning substitution is, however, conveyed by point estimates evaluated at the mean values of the cost shares. These elasticities are given in Table 5.3 using the formula (3.44) in Chapter 3. The asymptotic standard errors are in parentheses.

Looking at the table, several results stand out clearly. All cross elasticities of substitution are positive, implying that labor, capital, and energy are all substitutes in the production process. All three elasticities are different both from zero and from unity. Therefore, both the fixed-coefficient (Leontief) and Cobb-Douglas models would misrepresent the substitution possibilities among water supply inputs. These elasticities are calculated at the sample means. Attempts were also made to calculate the elasticities for each utility in the data. The same results were obtained. Labor and energy are substitutes in the table, but the estimated value is very low and insignificant. This is borne out in the sample too. Thirty-seven out of sixty utilities in the data show labor and energy substitution, others showing complementarity. Therefore, no strong substitution relation between labor and energy can be established at least for the utilities in the sample.

Table 5.3

Elasticities of Substitution Evaluated at Sample Means
(Asymptotic standard errors in parentheses)*

$\sigma_{LK} = 0.98166$ (0.01834)
$\sigma_{LE} = 0.03626$ (0.96246)
$\sigma_{KE} = 0.99538$ (0.97812)

*The elasticities of substitution (σ_{jq}) are estimated by:

$$\sigma_{jq} = (b_{jq} / S_j S_q) + 1$$

where S_j and S_q are the cost shares of inputs j and q . These elasticities are nonlinear functions of the estimated parameters, so that standard errors for their estimates cannot be calculated exactly. We therefore obtain approximate estimates of the standard errors under the assumption that the cost shares are constant and equal to the means of their estimated values. Under this assumption, we have asymptotically:

$$\text{var}(\sigma_{jq}) = \text{var}(b_{jq}) / S_j^2 S_q^2$$

However, it is apparent that the mean cost shares are not constant but stochastic, endogenous variables. Thus, a theoretically preferable procedure is to express the cost shares in terms of the regressors and the estimated parameters and to estimate the asymptotic standard errors by a linear Talor series expansion by including explicitly all variances and covariances; see Halvorsen (1977) and Moroney and Toeves (1977). However, there is no clear cut case for an advantage for either approach. In fact, the literature on the exact variance of products of random variables would lead to doubts over the reliability of either approach. One possible resolution of this issue can be derived from extensive sampling experiments.

It may be noted here that the elasticities in Tables 5.4, 5.5, and 5.10 are also nonlinear functions of the estimated parameters. The appropriate elasticities and their standard errors are evaluated at the sample means. However, as in the elasticities of substitution, exact distributions of the estimated standard errors are not known to us. For this reason, no conclusive statements about the significance of any elasticity estimates are made in this study.

It may be noted that the estimates of substitution among labor, capital, and energy are long run ex ante estimates. The received wisdom in the literature is that cross section analyses reflect long run adjustment possibilities, while the time series data yield short run estimates. The present research uses the cross section data. Thus, estimates of elasticities of substitution reflect the long run responses of input price changes. It may be true that, in water treatment, once a plant is established, little substitution is permitted among various inputs. Yet, it may not be unrealistic to suppose that, in the planning stage at least, the water supply manager is confronted with the possibility of choosing from a variety of input combinations.

In the short run, the input proportions of a large part of installed plant and equipment are insensitive to changes in the economic environment, having been chosen in response to the input prices in earlier periods. The response of the water planner is, therefore, restricted to adjustments at the "incremental" level, implemented through the purchase of new capital of contemporary vintage. In the long run, ceteris paribus, the constant decay of old capital ultimately results in the entire capital stock being replaced with "incremental" capital, and thereby acquiring input proportions that are in line with the changed economic environment.

Elasticities of substitutions are normalized input responses to changes in input prices. In order to evaluate the effect of price changes on the input usages, it is more appropriate to calculate own and cross price elasticities of input demands. Cross price elasticity η_{jq} is the change in the demand for input j due to a change in the price of input

q. If one is interested in a policy to affect the quantity of a specific input used in the production process by subsidizing input prices, cross price elasticities indicate which input prices it would be most effective to alter and in what direction they should be changed.

Own and cross price elasticities evaluated at the sample means are given in Table 5.4 utilizing the formulas (3.39) and (3.42) in Chapter 3. Each element in the row is the price elasticity of demand for the input after a price change of the input in the column. For example, L row refers to the price elasticity of labor with respect to a change in the price of labor, capital, and energy. The main diagonal elements are the own price elasticities of labor, capital, and energy, respectively.

All of the own price elasticities of input demand have the correct negative signs and are all significant. Own price elasticities of capital and energy are elastic, while the own price elasticity of labor is inelastic. The demand for capital is essentially unitary elastic. The demand for energy is more elastic relative to the demands for labor and capital. The cross price elasticities of input demands contain the same information as the elasticities of substitution, but unlike the elasticities of substitution, the cross price elasticities are not symmetric since they depend on the cost shares.

Suppose a 1% increase in the relative price of labor occurs. Then, because of an increase in labor cost, the water utility will be immediately induced to decrease the use of labor, e.g., by about 0.5% in this analysis. However, in order to produce a given level of outputs needed to serve consumers, the utility will be required to increase the use of either capital or energy. Since the cross price elasticity of

Table 5.4

Own and Cross Price Elasticities of Input Demand Evaluated at Sample Means (Asymptotic standard errors in parentheses)*

	W_L	W_K	W_E
L	-0.44611 (0.18244)	0.45380 (0.00848)	0.01672 (0.44494)
K	0.44317 (0.00838)	-1.23515 (0.17829)	0.44936 (0.44157)
E	0.00313 (0.08302)	0.08586 (0.08437)	-5.40505 (0.40606)

* Price elasticities of input demand are estimated by:

$$\eta_{jj} = s_j \sigma_{jj}$$

$$\eta_{jq} = s_q \sigma_{jq}$$

These elasticities are estimated at the mean cost shares. Asymptotic standard errors are estimated by:

$$\text{var}(\eta_{jj}) = \text{var}(b_{jj})/s_j^2$$

$$\text{var}(\eta_{jq}) = \text{var}(b_{jq})/s_j^2$$

energy with respect to the labor price is almost zero, this implies that substitution of labor toward energy is limited. Therefore, an alternative would be to build more labor-saving equipment and facilities to save the increased labor cost which may have hurt the operation of the utility. In fact, a 1% increase in the labor price will result in approximately 0.5% increase in the demand for capital.

Increases in capital cost have been a great problem for small water utilities. Small utilities do not have sufficient operating revenue and, as will be seen later, their unit cost of producing output is substantially higher than that of larger utilities. They also spend a larger portion of total cost on capital. Therefore, the substantial burden of capital cost has in the past induced many utilities to defer system maintenance and capital replacement. This is confirmed here. The own price elasticity of capital of -1.23515 suggests that for a 1% increase in the price of capital, the demand for capital will be decreased by more than 1%. This is exactly what has been happening in most water utilities. Most water utilities have been suffering from a substantial rise in capital costs, leading to deterioration of water systems and aging of capital facilities.

Since the cross price elasticity of energy with respect to the price of capital is almost zero, substitution of capital toward energy is negligible. As explained before, an increase in the price of labor does not have any substitution effect on energy either. The main reason is that energy is essential in the provision of water, despite the fact that it accounts for a small portion of the total cost. Energy is an indispensable input in the treatment process and in the transportation

of water to consumers. Thus, the use of energy would be least affected by a change in the price of capital or labor.

However, the cross price elasticity of labor with respect to the price of capital is positive, so labor is fairly easily substituted for capital. This implies that labor would be more susceptible than energy to a change in the price of capital. The utility will increase the usage of labor, for example, by making some workers perform more than one specialized task; i.e., the worker involved in supervision may also do some maintenance work. In fact, this is what has been happening in most water utilities in face of increases in capital and other costs in the provision of water. Operators, maintenance mechanics, and other workmen in water utilities, especially municipal utilities, have not been paid well in terms of wages, salaries, or employee benefits in relation to workers in other utilities and industries, but nevertheless they are heavily loaded with work; see AWWA report 1974, 1978. Therefore, we might conclude that high rental rates on capital increase the cost of using a given amount of capital, but the increase in capital cost is partially offset by substitution toward other inputs, primarily labor, as capital becomes more expensive.

Energy accounts for a relatively small portion of the utility's overall budget. However, in recent years, the price of energy relative to prices of other inputs has been increasing. Energy represents the fastest increasing cost category over the last few years; see Clark et al. (1979). We can determine the impact of a higher energy price on the utility's demand for inputs. The own price elasticity of energy is high and significant. For example, a 1% increase in the price of energy will

lead to a 5% decrease in the usage of energy. As will be seen below, this response seems to be overestimated for an "average" utility in our sample. But regardless of the magnitude of the elasticity, the immediate response of the utility to the higher energy price would be to cut down the usage of energy in the production of water. This can be done through energy conservation or through lower utilization of energy equipment. The curtailments in the use of energy, however, are likely to cause disruptions in production, unless either labor or capital is intensively used in the production process.

The cross price elasticity of labor with respect to the energy price is almost zero, thus labor substitution would not be possible. What could be done is to increase the usage of capital. In fact, a 1% increase in the price of energy would lead to a 0.5% increase in the usage of capital. This may not be possible in the immediate short run. But in the long run, the utility will adjust to the increased energy cost by utilizing existing technology to design and install production processes that have smaller energy requirements. The result is that production is more energy efficient but it requires a more intensive use of capital.

In fact, recent rising energy prices coupled with rising usage rates have caused the utility to become extremely energy conscious; see Clark et al. (1979). The utility has begun to optimize energy use by increasing system storage combined with increased off-peak pumping and the installation of more energy-efficient pumps, which require more capital. Thus, the ability of the utility to cope with the higher energy price is reduced to its ability to substitute other inputs for energy. The estimate of substitution between energy and capital is consistent

with results obtained for other industries such as manufacturing and electric utilities;⁴ see Griffin (1977), Halvorsen and Ford (1979), Pindyck (1979), and Field and Grebenstein (1980), among others.

The own and cross price elasticities of input demand are evaluated for a representative utility calculated at the sample means. But no utility in the sample may correspond (paradoxically) very closely to the hypothetical utility of the sample means. Therefore, the price elasticities are calculated for each utility in the sample. Most own and cross price elasticities are close in both magnitude and signs to the elasticities calculated at the sample means. However, own price elasticities for capital and energy averaged for all utilities in the sample are $\eta_{KK} = -0.527$ and $\eta_{EE} = -0.402$ and statistically significant at the conventional level of significance. Thus, the own price elasticities of capital and energy calculated for a representative utility at the sample means are markedly different in magnitude from the average elasticities.

⁴ It may be useful to note studies examining substitution possibilities among capital, labor, and energy. Greater agreement exists regarding the labor-energy relationship--at least most studies have found that the two inputs are substitutes, or essentially independent as in the case of Pindyck (1979, 1980). The estimated elasticities vary considerably, however. The widest area of agreement seems to be in the capital-labor substitutability, most of which are reasonably close to unity (with the exception of Griffin and Gregory (1976). Substitutability between energy and capital remains an unsettled question. Some researchers (for example, Berndt and Wood 1975) find these two inputs to be complements, while others find them to be substitutes. A number of attempts were made to reconcile the discrepancies in the energy and capital substitution estimate. Berndt (1976), Berndt and Wood (1979, 1981), Griffin (1981a, b), Field and Grebenstein (1980), Stapleton (1981), and Berndt, Morrison, and Watkins (1981), among others, note that the divergent results can arise due to: (1) the nature of data used, i.e., cross section versus time series; (2) the omission of material inputs when the remaining inputs are not in fact separable from materials; (3) the use of working versus physical capital; and (4) the specification of the underlying model; i.e., static versus dynamic model. It is not, however, a purpose of this research to attempt to address the difficult, touchy issue.

This is particularly so for energy. Therefore, we might say that the response of the energy price on its own demand is rather exaggerated for a representative utility. Wide difference in two price elasticities of energy reflects difference in the usage of energy across utilities. For small utilities, energy costs may represent a significant portion of their budget. Thus, an increased energy price relative to prices of other inputs will significantly affect the consumption of energy. This may not be true for large utilities. Large utilities may use less energy per unit of water treated than do small utilities due to more system storage and more flexibility in off-peak pumping.

An important policy implication of this discussion with respect to the own price elasticity of capital is the adverse impact of water quality regulation. The Safe Drinking Water Act discussed in Chapter 2, for example, may require water systems to install new technology which may in turn require a lot of additional capital in order to meet drinking water standards. As a result, utilities have to issue more bonds to finance the necessary plant and equipment. However, many utilities have reached the debt limit, thus prohibiting the issuance of more bonds; see Stevie and Clark (1982).

The ability of the very small utilities to finance capital is more severely limited than the large utilities. Thus, they may have to pay higher interest rates than the large utilities for financing. This implies that the opportunity cost of complying the Safe Drinking Water Act is higher for small utilities. The substantial increase in capital costs and opportunity cost would dampen the demand for capital required to install new technology, as shown by the estimate of the own price

elasticity of capital -1.23515. Further, if these costs are added to an already high cost for water supply, the economic burden to small system users would be high. Consequently, the successful implementation of the Safe Drinking Water Act might require some measures (e.g., in the form of a relief to interest rate) to help small utilities in alleviating the financial burden of acquiring capital funds needed to meet safe drinking water standards.

So far, the discussion has concentrated on the elasticities of substitution and the price elasticities of input demand. Before moving to the next topic, it is important to note that the input price elasticities are calculated under the assumption that outputs are held constant.⁵ They neglect the demand effect resulting from the change in the price of the output caused by increases in the price of any input. As a result, these elasticities probably underestimate the true input price elasticities where the effect of input price changes on output prices is taken into account. This is likely to be a serious underestimation for inputs comprising a large proportion of total costs. But for water

⁵ These elasticities are referred to as compensated (or Hicksian) elasticities of input demand. Allen (1938) has shown that ordinary (or Marshallian) elasticities of input demand are given by:

$$\eta_{jj}^* = S_j (\eta_{jj} + \eta)$$

$$\eta_{jq}^* = S_q (\eta_{jq} + \eta)$$

where η represents the elasticity of demand of the output which respect to its own price. The price elasticity of the final output enters because increases in the price of any input not only cause a substitution among inputs, but also cause the price of the output to rise, inducing a substitution away from the output. For the multiproduct firm, η can be interpreted as a weighed average of output price elasticities of different outputs.

utilities, it is unlikely that changes in the price of any input would have much impact on output price, in view of the fact that the price of water is set exogenously. Therefore, the output-constant (compensated) elasticities may not be serious underestimates of the true elasticities of input demand.⁶

Input Requirements Due to Changes in Demand for Water, Capacity Utilization, and Service Distance

The elasticities of input demand with respect to outputs and operating variables describe percentage changes in input requirements in response to a small change in the demand for water (residential and nonresidential), capacity utilization, and service distance. These elasticities are calculated at the sample means using the formulas (3.47) and (3.50) in Chapter 3 and given in Table 5.5. Each element in the row is the input requirement coefficient implied by a small percentage change in the variable in the column. For example, L refers to the labor requirement due to a change in residential demand for water, nonresidential demand for water, capacity utilization, and service distance.

A 1% increase in the residential demand for water would require an increase in about a 0.7% of the labor usage. The same percent increase in the residential demand for water requires a 0.8% increase in the

⁶ But, a large number of privately-owned utilities are self-supporting. For these utilities, an increase in cost must be recovered through an increase in the water rate. Then, there is a potential for a bias in the output-compensated elasticities of input demand. However, as mentioned before, more than 80% of the total supply of water is provided by municipal utilities which are the main concern in this study.

Table 5.5

Elasticities of Input Demands With Respect to Changes in Demand for Water (Residential and Nonresidential), Capacity Utilization, and Service Distance Evaluated at Sample Means (Asymptotic standard errors in parentheses)*

	Y_R	Y_N	Z_U	Z_D
L	0.72257 (0.42637)	0.21021 (0.27984)	-0.07175 (0.59938)	0.15767 (0.28119)
K	0.28144 (0.43123)	0.29055 (0.28333)	-0.03796 (0.60415)	0.46493 (0.28555)
E	0.80315 (0.45515)	0.33694 (0.30208)	0.22115 (0.62260)	-0.40342 (0.09223)

*These elasticities evaluated at the sample means of the data are obtained by:

$$\partial \ln X_j / \partial \ln Y_i = d_{ij} / S_j + a_i$$

$$\partial \ln X_j / \partial \ln Z_k = f_{jk} / S_j + c_k$$

Since the capital share was deleted in the estimation procedure, standard errors of the estimated elasticities corresponding to the K row are calculated under the assumption that:

$$\text{cov}(d_{iK}, a_i) = 0 \quad (i = R, N)$$

$$\text{cov}(f_{Kk}, c_k) = 0 \quad (k = U, D)$$

energy requirement. This suggests the importance of operating inputs, labor and energy, in the operation of utility. In particular, this indicates that the provision of water is an energy-intensive process. The supplying of water to consumers requires the intensive use of energy in the treatment and transportation of water. Energy is also required to produce the chemicals and other materials in the operation of utility--the secondary energy requirements. This includes the manufacture and regeneration of activated carbon which is an energy-intensive process; see Cornell (1978). Furthermore, increased regulatory requirements, e.g., the Safe Drinking Water Act, may result in increased energy use.

A 1% increase in the demand for residential water will cause about a 0.3% increase in capital facilities, a relatively low percentage figure. The reason for the low capital requirement might be that the excess capacity in the existing system can be utilized to fulfill the requirement of system growth or expansion without increasing plant investment. Thus, the capital requirement or capacity needs in response to an increase in the demand for water (residential or nonresidential) would be smaller than would otherwise be needed when the excess capacity is not utilized.

Input requirements of a change in demand for water are different for residential and nonresidential customers. This is due to the fact that residential customers constitute a majority of water users. Specifically, the revenue share for residential water is usually higher than the revenue share for nonresidential water. Therefore, a 1% change in residential water corresponds to a larger change in the value of the final output than does a 1% change in nonresidential water. Conse-

quently, it would not be surprising to find that the input demand elasticities for residential water are higher than the corresponding elasticities for nonresidential water.

Suppose now that the utility expects a 1% increase in both residential and nonresidential demand for water. then the required input changes can be obtained by adding corresponding inputs in the Y_R and Y_N columns. Thus, the utility would require about a 1% ($0.72257 + 0.21021 = 0.93278$) increase in the usage of labor, about a 0.6% ($0.28144 + 0.29055 = 0.57199$) increase in the usage of capital, and a 1.2% ($0.80315 + 0.33694 = 1.18009$) increase in the usage of energy. Again, the capital requirement for a small increase in the demand for water is low, but the energy requirement is relatively high.

The Z_U column shows the change in input requirements due to a change in capacity utilization. Notice that the input requirement coefficients of labor and capital are negative. This means that as the utility increases the rate of capacity utilization, it would need less labor and capital in the production process. However, more energy is needed since the energy requirement coefficient is positive. The consequence is that variations in capacity utilization rates can have a significant impact on the input demands for water supply.

While the implication of the negative sign for labor requirement may not be obvious, the negative sign for capital requirement may be explained as follows. The capacity utilization rate is defined as the ratio of the average use of system to the rated system capacity in this study. Thus, capacity utilization can be increased either with an

water, either residential or nonresidential) at the given level of capacity or with a decrease in the use of the system at the given level of the system demand. With the demand for water and other factors held fixed, a decrease in the use of the system or capital can be explained by a smaller reserve capacity for peaking, for example.⁷

As can be seen in the table, an increase in service distance Z_D requires more labor and capital but less energy. This seemingly paradoxical result stems from the fact that input requirement coefficients measure the effect of the ceteris paribus change in service distance on input requirements, holding other factor constant. Thus, an increase in service distance without a concomitant increase in the demand for water (requiring more water treated and pumped) would not require more energy nor chemicals.⁸ But it would require more labor and capital, since an increase in service distance requires installation of more pipelines.

An interesting question is then how much additional labor, capital, and energy would be needed if the water utility wishes to serve additional customers, i.e., new (not existing) customers in a rapidly growing area. This would require an increase in service distance

⁷ Everson (1982) notes that increases in capacity utilization or decreases in maintenance of system may cause a piece of capital equipment to deteriorate rapidly. He then develops a dynamic input demand model of the interaction of utilization, maintenance, and investment decisions and tests the model on railroad freight cars.

⁸ According to Dick Stevie, it might be true for gravity systems but not for pumping. For pumping, an increase in the distance might require more energy. Distance directly affects the amount of KWH used. Therefore, it may be that, in the data, the measure of distance reflects the fact that larger service areas have more energy-efficient pumps.

followed by more water produced. Consider a water utility wishing to serve 1% more residential customers by increasing service distance by 1%. Then the utility would require an increase in the usage of labor by about 0.8% ($0.72257 + 0.15767 = 0.80240$), of capital by about 0.7% ($0.28144 + 0.46493 = 0.74637$), and of energy by about 0.4% ($0.80315 - 0.40342 = 0.39973$). Energy requirement of about 0.4% is consistent with the speculation of Clark et al. (1978) that as the distribution area (and thus service distance) increases, the amount of energy used to serve customers will also increase.

These results demonstrate that capacity utilization and service distance have substantial effects on input requirements. Thus, efforts to analyze the input demand without considering these variables would lead to serious bias in the estimates of input requirements. For example, the exclusion of capacity utilization would lead to overestimation of labor and capital requirements but underestimation of energy requirements. The omission of service distance in the demand for input would lead to underestimation of labor and capital requirements but overestimation of energy requirements.

The implications of the discussion in this section are clear. Even though the prices of all inputs may remain constant, the dependence upon various inputs can be altered by changes in the distribution of customer demands as well as changes in capacity utilization rates and service distance. Thus, a utility in forecasting future input requirements must not only consider future input prices but also the future customer demands and other operating characteristics of the utility.

So far in this discussion on input responses due to changes in input prices, outputs, and operating variables, no mention has been made of the impact on costs of an increase in the demand for water, capacity utilization, and service distance. This will be discussed in the next few sections.

5.3 SCALE ECONOMIES

The question of scale economies in water utilities is a much debated issue. Previous studies in the estimation of scale economies were not always satisfactory. This analysis presents new estimates of scale economies and compares them with the estimates in the previous studies.

Scale Economies for Residential and Nonresidential Water Supply

Scale economies are measured by the cost elasticity of output. A high value of the cost elasticity of output implies decreasing scale economies, while a low value implies increasing scale economies. Since we have two outputs, residential and nonresidential water supply, we have two corresponding cost elasticities. These elasticities can be derived by utilizing of equation (3.52) in Chapte 3.

The cost elasticity of residential water supply is estimated in the following way:

$$\begin{aligned}
 E_{CR} = & 0.53037 - 0.42487 \ln Y_R + 0.10086 \ln Y_N \\
 & + 0.08885 \ln W_L - 0.11238 \ln W_K + 0.02353 \ln W_E \\
 & - 0.59305 \ln Z_U + 0.44056 \ln Z_D
 \end{aligned}$$

The cost elasticity of nonresidential water supply is given by:

$$\begin{aligned} E_{CN} = & 0.25741 + 0.10086 \ln Y_R + 0.14868 \ln Y_N \\ & - 0.02182 \ln W_L + 0.01496 \ln W_K + 0.00686 \ln W_E \\ & + 0.13495 \ln Z_U - 0.32537 \ln Z_D \end{aligned}$$

The cost elasticities of both residential and nonresidential water supply depend not only upon the levels of water produced but also upon other variables. This relationship reflects significant nonhomothecities present in the water supply technology with respect to outputs, input prices, and operating variables. These are the same variables that affect the cost function. Both residential and nonresidential water are included in each cost elasticity equation due to the joint costs incurred in the provision of water to residential and nonresidential users. Thus, the effects of residential and nonresidential water on the costs and scale economies cannot be separated.

In the cost elasticity equations for residential and nonresidential water supply, positive values of the coefficients indicate that increases in the corresponding variables lead to lower scale economies, while negative values indicate that increases in the corresponding variable lead to higher scale economies. The cost elasticities of residential and nonresidential water supply show marked difference in the coefficients (in both signs and magnitude) of the relevant variables. This stems mainly from the fact that residential water supply constitutes a large portion of water supplied to consumers. The coefficients of input prices in the two cost elasticities are generally

small relative to the coefficients to outputs, capacity utilization, and service distance. This implies that the effects of relative price movements of labor, capital, and energy are relatively minor in the determination of scale economies for residential and nonresidential water supply. Thus, we might conclude that scale economies are by and large determined by the levels of outputs, capacity utilization, and service distance. Noticeable among these variables are capacity utilization and service distance.

The residential water variable Y_R has a negative coefficient in the cost elasticity of residential water supply. This means that, other factors remaining unchanged, an increase in the quantity of water treated for residential customers leads to an increase in the value of scale economies for residential water supply. The capacity utilization variables Z_U has a negative coefficient too. Thus, the effect of an increase in capacity utilization is to increase scale economies for residential water supply.

By contrast, the effects of an increase in service distance is decreased scale economies for residential water supply, as evidenced by the positive coefficient of Z_D . Therefore, with the same flow of water and other conditions constant, the utility with larger service distance would expect lower scale economies for residential water supply than the utility with smaller service distance. As a consequence, we might say that the treatment activity for residential water supply is likely to subject the utility to more scale economies, while the distribution activity as reflected by service distance tends to subject the utility to more diseconomies of scale.

It may be noted that the residential cost elasticity can be calculated for each utility. At the sample means of the variables, the cost elasticity is $a_R = 0.53037$. The elasticity is positive and does not exceed or even approach the value of 1, constant returns to scale. This suggests that there are substantial economies of scale in the provision of water to residential customers.

The cost elasticity of 0.53037 can be viewed as being the weighted sum of the output elasticities of input demands for residential water supply discussed in the previous section, with the weights being the estimated input shares at the point of approximation. As discussed previously, an increase in residential demand for water creates additional demands for labor, capital, and energy, causing an upward pressure on costs. The cost elasticity of 0.53037 implies that a 1% increase in residential water brings about a 0.53% increase in total cost. As shown in the Y_R column in Table 5.5, the 0.53% increase in cost can be seen to consist (approximately) of a 0.72% increase in labor usage, a 0.28% increase in capital usage, and a 0.80% increase in energy requirement.

Unlike residential water supply, the cost elasticity of nonresidential water tells us a little different story. As indicated before, the difference results mainly from the fact that residential water constitutes the majority of water supplied to consumers. For example, the coefficients for residential and nonresidential water usage Y_R and Y_N are positive in the cost elasticity of nonresidential water supply. Thus, the effect of more water produced, *ceteris paribus*, is to decrease nonresidential scale economies. This is in opposition to the case of

residential water supply. The sign of capacity utilization in the nonresidential cost elasticity is also positive. This is again in contrast to the residential water case, but it is partly attributable to smallness of nonresidential water relative to residential water.

Another striking fact about the nonresidential cost elasticity is the negative sign of the service distance variable. This suggests the effect of an increase in service distance in order to serve additional nonresidential customers is an increase in nonresidential scale economies, a result which is in opposition to that of residential water supply. Therefore, the provision of water to nonresidential customers is characterized by cost diseconomies in treatment but economies in the distribution system.

The cost elasticity of nonresidential water supply can be calculated for each utility. At the sample means, the elasticity is $a_N = 0.25741$. Thus, a 1% increase in nonresidential water supply would lead to about 0.26% increase in total cost. This implies that, as in residential water supply, there are marked economies of scale in the provision of water to nonresidential customers.⁹ As shown in the Y_N column in Table 5.5, the 0.26% increase in cost brought about by a 1% increase in nonresidential demand for water can be seen to consist (approximately) of a 0.21% increase in labor usage, a 0.29% increase in

⁹ In Section 3.5, it was noted that the cost elasticity of the individual output overestimates the output-specific scale economies. Therefore, it is expected that the degrees of actual residential and nonresidential scale economies are smaller than the estimates derived here. Yet, they do not change the conclusions drawn with respect to the presence of scale economies in both residential and nonresidential water supply. For example, assuming 65% and 35% of residential and nonresidential cost shares (which are not realistic), the output-specific scale economies for residential and nonresidential water usage are 0.81677 and 0.73619, respectively.

capital usage, and a 0.34% increase in energy requirement.

The cost elasticity of nonresidential water supply is lower than the cost elasticity of residential water supply, implying greater scale economies of nonresidential water supply in comparison to residential water supply. This indicates that the average cost of supplying water to nonresidential customers declines more rapidly than to residential customers as a group. This is in accord with prior expectation, the reason being that a given amount of water involves more service lines and more connections for residential use than for business and industrial uses. This is consistent with the result obtained earlier on the difference in the input requirements with respect to increases in demand for residential and nonresidential water. For the same increases in both residential and nonresidential usage with input prices and other conditions fixed, the input requirements for residential usage are higher than those for nonresidential water usage. This would make the cost elasticity of residential water supply higher than the cost elasticity of nonresidential water supply. By implication, nonresidential users of water are more economical to serve than residential users as a group.

Previous studies were not able to confirm this result. For example, an EPA study (1978) estimated the cost functions for residential and nonresidential water supply separately, which amounts to assuming that the provision of residential and nonresidential water supply are nonjoint. A simple regression was run whose arguments included outputs and service distance. This study obtained a rather paradoxical result of the cost elasticity of nonresidential water supply

being greater than the cost elasticity of residential water supply. Many factors can be attributed to this result, but primary causes would be the misspecification of technology and the use of incorrect functional form.

Most studies in water utilities have not broken water supply into customer classes in the calculation of scale economies. Therefore, it will be of interest to calculate the overall scale economies. The overall scale economies can be obtained by adding the two cost elasticities of outputs, residential and nonresidential water supply.

Overall Scale Economies

Assuming both residential and nonresidential water supply increase in the same proportion, i.e., letting $d\ln Y_R = d\ln Y_N = d\ln Y$, we can express the overall cost elasticity as:

$$E_{CY} = 0.78778 - 0.32401 \ln Y_R + 0.24954 \ln Y_N + 0.06703 \ln W_L \\ - 0.09742 \ln W_K + 0.03039 \ln W_E - 0.45810 \ln Z_U + 0.11519 \ln Z_D$$

Without the breakdown of water supply, i.e., $Y_R = Y_N = Y$, the overall cost elasticity can be seen as a counterpart to a measure of scale economies for a single output.

Most of the signs (not the magnitude) of the coefficients are the same as the signs of the residential cost elasticity, since the effects of the residential cost elasticity override the effects of the nonresidential cost elasticity. Therefore, the same interpretation of the coefficients can be offered to the overall scale economies as that given to the residential scale economies.

An important finding is that an increase in the treatment activity decreases the overall cost elasticity of output, as shown by the negative coefficients of Y and Z_U .¹⁰ This implies that an increase in the treatment of water, *ceteris paribus*, is likely to make the water supply firm more subject to economies of scale. By contrast, an increase in service distance increases costs (as will be shown in the next section), but such increase also serves to increase the cost elasticity of output, as shown by the positive coefficient of Z_D .¹¹ This means that increases in service distance will tend to make the firm more subject to diseconomies of scale, *ceteris paribus*.

Therefore, with input prices held constant, the overall scale economies can be determined by comparing economies in treatment with diseconomies in distribution as measured by service distance. The apparent finding of substantial diseconomies with respect to distance indicates that there are rather definite limitations to the economic size of water supply areas. A critical factor will then be cost tradeoffs involving treatment economies with diseconomies of distribution. If distribution costs outweigh facilities costs and the assumed economies of scale associated with water treatment, then decentralization of production to many locations would be economical. Should the opposite be true, centralization of production is preferable (Becksenstein 1975).

¹⁰ This is consistent with the inverse relationship between unit cost and capacity utilization found by Hines (1969).

¹¹ This is in full agreement with a positive relationship found in the EPA study (1978) to exist between distance and average cost.

Overall scale economies vary across utilities. The overall scale economies at the sample means are the sum of the cost elasticities of residential and nonresidential water supply at the sample means, i.e., $0.53037 + 0.25741 = 0.78778$. The elasticity is positive and less than one. This implies that the underlying technology of water supply exhibits considerable economies of scale. Thus, earlier results concerning the persistence of economies of scale in water utilities are substantiated in our finding. The important implication of this finding concerns the nature of production costs. In fact, for the average utility, a 1% increase in both residential and nonresidential water supply results approximately in an increase of about 0.8% in total cost. Given this result, we can conclude that the average utility chosen in the sample is operating in the region of declining average costs.

Since scale economies are much greater for nonresidential supply of water as opposed to residential supply, we may conclude that economies of scale in water supply are mainly determined by nonresidential demands of water. The EPA study (1978) as noted before, however, made a contrary conclusion. This is due to a paradoxical result of the cost elasticity of nonresidential supply of water being greater than that of residential supply.

The overall scale economies are calculated for each utility and the majority of utilities show economies of scale. These elasticities are presented in Table 5.6 for selected utilities. The overall scale economies increase more or less with an increase in output (total water treated for residential and nonresidential usage) and there is some evidence that economies of scale exhaust at larger scales of operation.

Table 5.6
Overall Scale Economies for Selected Utilities

Utility	Total Water Treated (MGD)	E_{CY}
Boring	0.044	0.11290
Algonquin	0.490	0.55403
Dallas, WCID	1.093	0.44412
Killen	4.482	0.32939
Falls Church	14.611	0.53481
East Jefferson	33.896	0.58266
New Haven	58.367	0.71444
Kansas City	95.992	0.95569
Seattle	151.753	0.99457
Philadelphia	372.784	1.93172

The minimum efficient scale of operation--the smallest output level for which the cost elasticity of output is unity, or equivalently the smallest output at which average cost reaches its minimum value--appears to be about 100 MGD. Assuming a 60% load factor (which is the sample average), this corresponds to a utility with 167 MGD of capacity. If the system becomes too large, the water utility can experience diseconomies of scale, as evidenced by the Philadelphia water utility in the table.

An important policy implication of this result pertains to the regionalization or mergers of water utilities discussed in Chapter 2. Underlying the recommendation for the regionalization in water utilities is the belief that this will allow utilities to achieve economies of scale which would otherwise be unattainable. Since this study found most utilities enjoying substantial economies of scale, this would provide an economic justification for regionalization of utilities, a result which is in conflict with Mann and Mikesell (1976). For a given regulated rate structure, a larger scale of operation associated with regionalization should yield lower costs and higher profits for regionalized utilities. However, large utilities (such as Seattle and Philadelphia) would have little incentives to merge with other utilities, since they have already fully exploited existing economies of scale. Mergers by these utilities would needlessly reduce the scope for potential competition in the water supply industry.¹² The industry not only becomes more concentrated, but

¹² However, it might be argued that mergers among large utilities would still be desirable if it would be demonstrated that large utilities innovate at a significantly higher rate than small utilities.

large utilities also become significantly larger. Mergers among small utilities can result in direct cost savings and the creation of additional utilities of efficient size, thereby improving the prospect for more real competition among equally efficient utilities.

The estimates of the overall scale economies obtained in this analysis are significantly different from those of the previous investigations for the following reasons. First, almost all of the previous studies use the functional breakdown of the cost function into system components -- acquisition, treatment, distribution, transportation, and overhead. The breakdown of the cost function is based on the assumption that the water supply activities are separable and each activity has its own unique cost function, implying absence of economic considerations linking the treatment and distribution activities. The breakdown of systems is thus unnecessarily restrictive and involves some arbitrariness. In fact, the process of acquisition, treatment, distribution, and transportation of water can be best characterized as a joint activity. Our estimates of the overall scale economies are for the utility's complete production activity and are not based on the restrictive disaggregation of system components.

Second, even with the breakdown of system components, the previous studies employ restrictive functional forms. As indicated in an earlier chapter, because of the extremely simplified functional forms such as Cobb-Douglas, scale economies are assumed to be constant for all levels of output. Unfortunately, this is contrary to what the theory of the firm tells us. The theory of the firms says that scale economies may vary not only with the levels of output but also with other important

variables such as input prices. Therefore, there are substantial errors in the estimates of scale economies in the existing studies. As a consequence, there has been significant disagreement as to the range over which scale economies, if any, persist in water supply. The estimates of scale economies produced via this research are calculated using the flexible functional form. And, the values of scale economies vary with the levels of outputs and with other variables.

Third, previous studies have not directly incorporated capacity utilization and service distance into the estimation of scale economies. This study shows that the effect of any increase (decrease) in capacity utilization is certain to cause a decrease (increase) in the cost elasticity of output. Thus, if the utility experiences a decrease in the capacity utilization rate, perhaps because of capital expansion involving lumpy expenditures on increasingly larger units, the inability to account for capacity utilization would underestimate the true scale economies. By contrast, the effect of an increase in service distance is to increase the cost elasticity of output. As a result, previous studies that ignore service distance in the estimation of the cost function tend to underestimate the true scale economies.

Thus, the failure to account for change in capacity utilization and service distance is a serious error that leads to biased estimates of scale economies. The net effect of excluding both capacity utilization and service distance on the estimates of scale economies will vary across utilities. The methodology employed here directly incorporates the capacity utilization and service distance variables in the estimation of the cost function. Thus, the estimates of scale economies take into

account the effects of capacity utilization and service distance. An important implication of this result is that regulatory decisions and policies can have significant impact on scale economies by directly affecting capacity utilization rates and service distance. It also implies that regionalization attempts which affect service distance, for example, by interconnecting various hydraulic networks have to be analyzed with greater care since they might offset the scale economies achieved in water treatment.

Scale Economies in Relation to Spatial Characteristics

A number of studies suggest that economies assumed in the treatment of water may be sacrificed by diseconomies in the distribution of water, which is directly affected by the characteristics of the area and neighborhoods being served; see Dajani (1973) and Gemmell (1976). As a result, these studies conjecture that water utilities may be characterized by diseconomies of scale; see Clark (1979). They further point out that there may be an efficient size of water service area with respect to distance; see Clark and Stevie (1981).

An investigation of this conjecture means to determine whether diseconomies of service distance exceed economies in water treatment. If so, the net effect of an increase in service distance would lead to overall diseconomies of scale in water utilities. Otherwise, we might conclude that the overall water utility activities are characterized by economies of scale despite diseconomies present in the distribution of water.

In most utilities, service area and distance are closely related to

the size of utility. Large utilities tend to have large service areas covering large distances. Thus, the effect of an increase in service distance cannot be separated from the effect of an increase in the size of a utility. Large utilities enjoy substantial economies of scale in water treatment relative to small utilities. But, at the same time, these utilities suffer from substantial diseconomies due to the size of the area being served. To see the relationship among the system size, service distance, and scale economies, a column with service distance is added to Table 5.6 for the selected utilities. The new table is given in Table 5.7. The general pattern in the table is that utility with large distances tend to have the high cost elasticity. Scale economies depend not only on the size of water utility, i.e., total water treated, but on the service distance as well. Thus, the inclusion of service distance in the calculation of scale economies only reduces the extent of scale economies, but does not change the earlier conclusion about the existence of economies of scale in water utilities. One plausible explanation is that economies prevailing in the treatment of water are so dominant that there are no offsetting diseconomies in the distribution of water. Therefore, this finding does not support Dajani and Gemmel's conjecture that when distribution costs are added, diseconomies of scale may occur in water supply.

An interesting issue for investigation is to find the efficient size with respect to service distance. Since service distance is closely related to the size of utility, finding the minimum distance is equivalent to finding the utility with the minimum efficient scale of operation. Earlier, the efficient scale of operation was found to be

Table 5.7
Relationship Among System Size, Service Distance and Scale Economies

Utility	Total Water Treated (MGD)	E _{CY}	Service Distance (Miles)
Boring	0.044	0.11290	7.0
Algonquin	0.490	0.55403	6.0
Dallas, WCID	1.093	0.44412	65.0
Killen	4.482	0.32939	128.0
Falls Church	14.611	0.53481	349.0
East Jefferson	33.896	0.58266	781.6
New Haven	58.367	0.71444	1266.0
Kansas City	95.992	0.95569	1912.0
Seattle	151.753	0.99457	1547.0
Philadelphia	372.784	1.93172	3254.0

about 100 MGD. Looking at Table 5.7, the minimum distance corresponding to the system with 100 MDG appears to be about 1500 miles. If the utility service distance reaches beyond this minimum distance, the utility will exhibit diseconomies of scale due to diseconomies in the distribution of water outweighing economies in the treatment of water.

5.4 RELATIONSHIP BETWEEN OPERATING VARIABLES AND COSTS

The effects of an increase in capacity utilization or service distance on input requirements and scale economies were discussed in the last section. In this section, we will discuss the impact of an increase in capacity utilization or service distance on total cost.

Capacity Utilization

The elasticity of cost with respect to capacity utilization measures the percentage change in total cost due to a small percentage change in capacity utilization. Making use of equation (3.58a) in Chapter 3, the cost elasticity with respect to capacity utilization is given by:

$$\begin{aligned} E_{CU} = & -0.03123 + 0.38587 \ln Z_U + 0.44037 \ln Z_D \\ & - 0.59305 \ln Y_R + 0.13495 \ln Y_N - 0.01873 \ln W_L \\ & - 0.00304 \ln W_K + 0.02177 \ln W_E \end{aligned}$$

Given the nonhomothetic technology of water supply, the cost elasticity of capacity utilization depends not only upon the rate of capacity utilization but also upon the actual levels of water (both

residential and nonresidential) treated and other variables. The sign of the capacity utilization variable is positive, imply that, *ceteris paribus*, an increase in capacity utilization would lead to an increase in the value of the cost elasticity of capacity utilization. The signs of the input price variables are rather ambiguous but they reflect the response of the input prices on the cost elasticity. The signs on residential and nonresidential water usage are different. An increase in residential water usage increases the elasticity. This reflects the fact that residential water supply constitutes a large portion of water supplied to consumers.

The cost elasticity of capacity utilization varies with utilities in the sample. But we can evaluate it for a representative utility calculated at the mean levels of the variables. The cost elasticity of capacity utilization at the sample means is $a_U = -0.03123$. Thus, a 10% increase in capacity utilization rate would result approximately in a decrease of 0.3% in total cost. This suggests some economies of capacity utilization. The economies are possible due to decreased input requirements for labor and capital relative to those of energy, as seen in Section 5.2.

The cost elasticity of capacity utilization was also calculated for each utility in the sample. The majority of utilities show the negative values averaging -0.008. Thus, an increase in capacity utilization results in an increased cost saving.

Marginal cost of capacity utilization is given by:

$$\frac{\partial C}{\partial Z_U} = \frac{\hat{C}}{Z_U} E_{CU}$$

where $\hat{C}$ is the fitted total cost. Marginal cost of capacity utilization depends upon the same variables that affect the cost elasticity of capacity utilization. At the sample means ($C = \$224.610814$ in thousands of dollars, $Z_U = 0.56282$), marginal cost of capacity utilization is $\$124.38347$ in thousands of dollars. Thus, a one unit (for example 1%) increase in capacity utilization can result in an additional saving of about $\$124$ in thousands of dollars in total cost.

Change in capacity utilization affects the input requirements and the cost structure of water supply. Thus, the inability to account for capacity utilization would lead to some spurious inferences about the structure of water supply. Despite its importance in the estimation of cost of water supply, capacity utilization has not been well recognized in previous studies. Early works include Konig (1967), who estimated average cost for complete treatment for different capacity utilization ratios. Hines (1969) estimated the relationship between percent capacity utilization rate and cost. However, the importance of Konig and Hines does not lie in the statistical and methodological merits but in the fact that it is one of the first attempts to include capacity utilization in the estimation of cost of water supply.

A more recent study is Mann and Mikesell (1979). They used a simple regression equation relating water cost to water produced and capacity utilization as measured by load factor. They obtained the positive relationship between capacity utilization and total cost, which implies that as the water utility utilizes the capacity more fully, the cost will

increase. This is contrary both to expectation and to our result. A number of factors can be attributed to this apparent paradox, among which misspecification of the technology of water supply is included.

Service Distance

The elasticity of cost with respect to service distance measures the percentage response in total cost due to a small percentage change in the distance being served by a water utility. Using equation (3.58b) in Chapter 3, this cost elasticity can be given by:

$$\begin{aligned} E_{CD} = & 0.24798 + 0.44037 \ln Z_U - 0.08081 \ln Z_D \\ & + 0.44056 \ln Y_R - 0.32537 \ln Y_N - 0.04175 \ln W_L \\ & + 0.09794 \ln W_K - 0.05619 \ln W_E \end{aligned}$$

The cost elasticity of service distance is a function not only of service distance but also of other variables. This is due to a significant nonhomothetic relationship present in water supply with respect to outputs, input prices, capacity utilization, and service distance. The coefficients of the input price variables are rather small. An increase in residential water supply increases the cost elasticity of distance, while an increase in nonresidential water supply decreases the elasticity.

The cost elasticity of service distance varies across utilities. At the sample means, the cost elasticity is $a_D = 0.24798$. Thus, a 1% increase in service distance would increase total cost by about 0.3%. This is because an increase in distance requires installation of more

pipelines and employment of more labor (as discussed in Section 5.2), resulting in an increase in costs.

The cost elasticity of service distance is evaluated for selected utilities. This elasticity is tabulated in Table 5.8, along with the corresponding marginal cost measured in \$1000/miles. The marginal cost is an additional cost incurred due to one additional unit increase in service distance holding quantities of water and other variables constant. Most utilities have a cost elasticity of service distance less than 1. The corresponding marginal cost ranges from \$0.7341/mile to \$2.46517/mile. There is no consistent pattern in the estimates of the cost elasticity of service distance and the corresponding marginal cost, however.

The cost elasticity of service distance measures the impact of service distance on cost. But an increase in service distance is usually accompanied by an increase in customer demand for water which will cause additional demands for inputs. Thus, these considerations have to be given in the evaluation of the cost of elasticity of service distance.

Suppose a water utility wishes to serve 1% more residential customers by increasing service distance by 1%. Then, the required input changes for these concurrent activities are calculated as discussed in Section 5.2. The effects on costs can be measured by adding the cost elasticities of residential water supply and service distance. At the sample means, the sum of the two elasticities is $a_R + a_D = 0.77835$ ($0.53037 + 0.24798$). Therefore, a 1% increase in demand for residential water with a 1% increase in service distance would bring about a 0.8% increase in total cost.

Table 5.8

Cost Elasticity of Service Distance and
Corresponding Marginal Cost for Selected Utilities

Utility	Total Water Treated (MGD)	Service Distance (Miles)	E_{CD}	MC_D
Cranby	0.216	5	0.26336	1.20848
Rifle	0.753	22	0.21563	0.73411
Georgetown	1.515	42	0.68933	1.12805
Rockville	4.110	132	0.33426	0.94183
Sun Valley	0.688	200	0.09247	0.19878
East Jefferson	33.896	781	0.50329	1.35882
New Haven	56.164	1266	0.36913	1.90452
Elizabeth	117.258	1790	0.36398	2.46517
Dallas	179.452	3208	0.29799	1.05372
Cincinnati	135.540	3785	0.19713	0.77122

Note: E_{CD} = the elasticity of cost with respect to service distance

MC_D = corresponding marginal cost measured in \$1000/mile

Coase (1947) was one of the earlier writers who pointed out that the cost of supplying water to consumers depends on quantity supplied and demand. He assumes that the marginal costs of supply to each consumer rise with an increase in the distance from the treatment plant, but it is constant at a given distance from the production source. As a result, the shape of the cost depends on the quantity consumed by consumers at various distances from the plant. Recently, Clark (1979) stressed the importance of the service distance variable in the estimation of the cost function and the implication on scale economies.

Overall, the findings of this section indicate that water supply costs are quite sensitive to operating conditions such as capacity utilization and service distance. In the next section, the results of other aspects of water supply operation are presented.

5.5 MARGINAL COSTS AND OUTPUT SUPPLY

This section presents estimates of marginal costs of residential and nonresidential water usage and the corresponding elasticities with respect to the variables affecting them. It then examines the transformation function of water supply.

Marginal Costs of Water

At a given point in time, the marginal costs of supply of water vary over many dimensions, among which are water treated, customer classes, and distance. Given this multidimensional nature of marginal costs, there have been few attempts at estimating marginal costs of water supply. Unfortunately, these attempts have not been successful for the

following reasons. First, the regression equations are crudely ad hoc with no theoretical considerations on the specification of technology; see Turvey (1976), Hanke (1978), and Bruggink (1982). Second, the definitions used in these research endeavors involve conceptual and pragmatic problems,¹³ often obscuring appropriate determination of marginal costs. The use of the flexible specification of technology eliminates many problems encountered in the previous research.

For the translog multiproduct cost function, marginal costs for residential and nonresidential water usage are derived from their respective cost elasticities E_{CR} and E_{EN} . Marginal cost of residential water usage is given by:

$$MC_R = \frac{\hat{C}}{Y_R} E_{CR}$$

where

$\hat{C}$ is the fitted cost value, and

$$\begin{aligned} E_{CR} = 0.43037 = & 0.42487 \ln Y_R + 0.10086 \ln Y_N \\ & + 0.08885 \ln W_L - 0.11238 \ln W_K + 0.02353 \ln W_E \\ & - 0.59305 \ln Z_U + 0.44056 \ln Z_D \end{aligned}$$

The marginal cost of nonresidential water usage is given by:

$$MC_N = \frac{\hat{C}}{Y_N} E_{CN}$$

¹³ See Mann (1981) for various definitions of marginal costs used in the water supply cost literature.

where

$$\begin{aligned}
 E_{CN} = & 0.25741 + 0.10086 \ln Y_R + 0.14868 \ln Y_R \\
 & - 0.02182 \ln W_L + 0.01496 \ln W_K + 0.00686 \ln W_E \\
 & + 0.13495 \ln Z_U - 0.32537 \ln Z_D
 \end{aligned}$$

Therefore, the marginal cost of supplying water varies across customer classes such as residential and nonresidential users. And, marginal cost of each customer class depends upon quantity consumed as well as other important variables, including service distance. These are the same variables that affect scale economies. Therefore, the calculation of marginal costs directly incorporates the variation in costs with respect to water produced, customer class, and distance.

Thus, even if two consumers--residential and nonresidential users--are located in the same service zone, the cost conditions would be different because of differences in other variables such as quantity consumed. Or, two residential and nonresidential users consuming the same amount of water would experience different cost conditions because of the differences in service distance. Furthermore, two residential customers would incur different cost characteristics, because of differences in quantity consumed.

It should be recognized that marginal costs of residential and nonresidential water usage can be calculated for each utility in the sample. To facilitate comparison with the prices of water to be discussed in the next section, marginal costs are expressed in terms of cents per

thousand gallons of water per year. This is done by converting the dimension of water flow from million gallons per day (MGD) to million gallons per year. At the sample means of the relevant variables,¹⁴ marginal costs of residential and nonresidential water usage are given by $MC_R = 16.06\text{¢}/1000$ gallons and $MC_N = 14.38\text{¢}/1000$ gallons, respectively. Thus, marginal cost of supplying water to residential users is higher than the marginal cost of supplying water to nonresidential users. This is consistent with the previous result that the cost elasticity of residential water supply is higher than the cost elasticity of nonresidential water supply.

Marginal Cost Elasticities

It will be useful to determine how marginal costs would vary due to changes in outputs, input prices, and operating variables. To do this, we can utilize the marginal cost elasticities given in Section 3.7. These elasticities are evaluated at the means of the variables and presented in Table 5.9. Each element of the row in the table is the marginal cost elasticity of output implied by a 1% change in the variable in the column. The overall effect on marginal cost of an increase in the variable in the column can be obtained by adding the corresponding MC_R and MC_N rows.

Looking first at the marginal cost elasticities with respect to the output variables Y_R and Y_N , we find that the "own" marginal cost

¹⁴This corresponds to $C = \$2241.610814/1000$ gallons, $Y_R = 7402.47740$ million gallons per year ($20.28076\text{MGD} \times 365$), and $Y_N = 4012.01795$ million gallons per year ($10.99183\text{MGD} \times 365$).

Table 5.9
Marginal Cost Elasticities With Respect to Outputs, Input Prices, and
Operating Variables Evaluated at Sample Means
(Asymptotic Standard Errors in Parentheses)*

	Y_R	Y_N	W_L	W_K	W_E	Z_U	Z_D
MC_R	-1.27071 (1.23766)	0.44758 (0.69736)	0.62981 (0.08025)	0.23956 (0.08788)	0.13065 (0.05411)	-1.14941 (0.87852)	1.07865 (0.64523)
MC_N	0.92220 (1.43685)	-0.16499 (1.06989)	0.37751 (0.11216)	0.50957 (0.12167)	0.11291 (0.07552)	0.49303 (0.92549)	-1.01604 (0.99522)

*For the calculation of these elasticities, see Section 3.7. Asymptotic standard errors are obtained by estimating the cost elasticities of output at the fitted mean values.

elasticities are all negative, as shown by $\partial \ln MC_R / \partial \ln Y_R = -1.27071$ and $\partial \ln MC_N / \partial \ln Y_N = -0.16499$, respectively. These results imply that increases in residential and nonresidential water supply are likely to decrease their own marginal costs. However, the marginal cost elasticity of residential water supply is greater than that of nonresidential water supply. This indicates that by increasing the quantities of residential and nonresidential water supply by an equal percentage, a greater decrease in marginal cost will be associated with residential water supply. We might think that this is contrary to the fact that the cost elasticity of residential water supply is greater than that of nonresidential water supply. But, as was demonstrated in Section 3.7, this will be resolved by knowing that there is no correspondence between marginal costs and scale economies.

The "cross" marginal cost elasticities of residential and nonresidential water supply are positive, as shown by $\partial \ln MC_N / \partial \ln Y_R = 0.92229$ and $\partial \ln MC_R / \partial \ln Y_N = 0.44758$, respectively. By implication, an increase in the quantity of residential water usage causes an increase in the marginal cost of nonresidential water supply. Also, an increase in nonresidential water usage leads to an increase the marginal cost of residential water supply. However, we would normally expect that the marginal cost of an output decreases when the quantities of other outputs increase. This is particularly so in the presence of economies of joint production; see Sakai (1974). Since we obtain the opposite result, this indicates the absence of economies of joint production associated with

residential and nonresidential water supply (for the average utility).¹⁵ As will be shown later, this arises because as the utility specializes in the production of either residential or nonresidential water, there is a substantial amount of saving associated with it by producing separately.

The marginal cost elasticities with respect to input prices are all positive, as shown by the W_L , W_K , and W_E columns in Table 5.9. Thus, the marginal cost of an output tends to increase when the prices of inputs increase. This is normally what we would expect. In general, an increase in the price of labor has a higher effect on the marginal cost of residential water supply than do the increase in the price of other inputs. An increase in the price of capital has a higher effect on the marginal cost of nonresidential water supply.

As in the discussion of the input demands in Section 5.2, we notice a substantial effect of operating variables in the operation of water utilities as shown by the Z_U and Z_D columns in Table 5.9. An increase in capacity utilization has a negative effect on the marginal cost of residential water supply but a positive effect on the marginal cost of nonresidential water supply. Conversely, an increase in service distance has a positive effect on the marginal cost of residential water supply but a negative effect on the marginal cost of nonresidential water supply. This agrees with the effects of capacity utilization and service distance on the cost elasticities of residential and nonresidential water supply.

¹⁵ This is also confirmed by the condition $a_{RN} + a_{R N} > 0$ given in (3.36) in Section 3.3.

An increase in capacity utilization has a negative effect on the aggregate marginal cost. But an increase in service distance has a positive effect on the aggregate marginal cost. Again, this is consistent with the effects of capacity utilization and service distance on the overall scale economies discussed in Section 5.3. An important implication of this finding is that marginal cost of supplying water can be significantly affected by changes in capacity utilization and service distance.

Transformation Function of Water Supply

The marginal rate of transformation between residential and nonresidential water supply is approximated by the ratio of the respective marginal costs. The ratio of the marginal costs of residential and nonresidential water supply evaluated at the sample means is given by $MC_R/MC_N = 1.11674$. The positive marginal rate of transformation indicates that opportunity costs are positive and substitution is necessary between residential and nonresidential water supply.

To check whether the firm incurs increasing or decreasing opportunity costs as it specializes in the production of either residential or nonresidential water supply, the curvature of the transformation curve is computed using the formula (3.75) in Chapter 3. At the sample means, the curvature is positive, indicating the convexity of the transformation frontier. This implies that the relative costs are decreasing due to further specialization in either residential or nonresidential water supply. This further indicates that the water supply firm would gain from

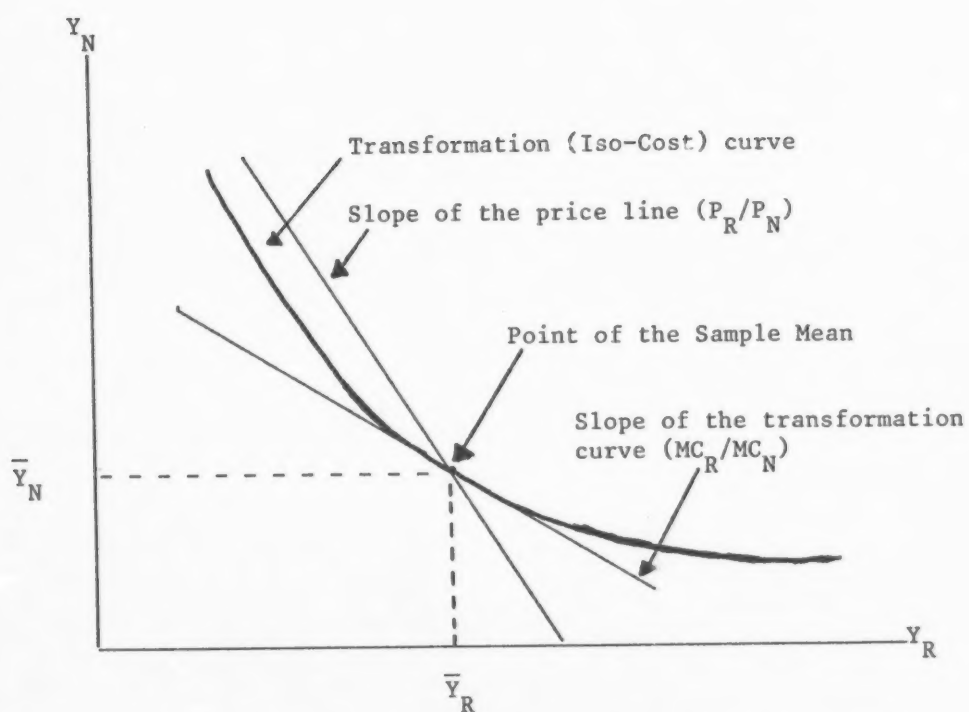
complete specialization in the production of either residential or nonresidential water supply.

The question is now whether, given the price ratio of residential and nonresidential water supply, the firm achieves the full competitive equilibrium. To investigate this question, we need to compare actual price ratio with the ratio of marginal costs of two outputs. At the sample means of outputs, the price ratio is larger than the marginal cost ratio. Hence the firm would prefer to increase nonresidential water supply relative to residential water supply at the existing prices. Figure 3 shows the geometrical demonstration of the above result where the graph is self-explanatory. It may be noted, however, that the transformation curve is evaluated at the given (sample mean) capacity utilization rate. Therefore, the curve shifts outward or inward for the different rates of capacity utilization.

5.6 PRICING OF WATER AND CROSS-SUBSIDIZATION AMONG CUSTOMER CLASSES

In the previous section, estimates of marginal costs and their relation to the transformation function of water supply were analyzed. This section examines pricing rules in relation to marginal costs of water supply. Actual prices of residential and nonresidential usage of water are first compared to their respective marginal costs. An alternative, second-best pricing rule is then discussed. The question of cross-subsidization among user classes is also discussed. In addition, the issue of whether water is underpriced or overpriced is analyzed.

Figure 3
Transformation Frontier



Actual Prices-Marginal Cost Ratios

Marginal costs in general reflect opportunity or real costs of obtaining water. These costs reflect the immediate and near term future cost consequences of input usage rather than the historical cost consequence of consumption.¹⁶ Thus, the estimation of marginal costs is "forward looking" and involves the forecasting of future cost-output streams. These cost-output projections require information on several variables, including technology, input behavior, and price elasticity of demand for water. As a result, costs and prices as determined under a regime of historical cost will differ from the marginal costs and prices based on marginal costs.

The common practice in the rate design in water supplies is based on historical or embedded average costs instead of marginal costs. It is therefore of interest to calculate the actual prices of water based upon the historical cost figures for comparison to the marginal costs calculated in this study. However, because water utilities do not keep good records on the breakdown of sales into customer classes, there is no direct information on the prices of residential and nonresidential water usage from the data set used in this study. Fortunately, the American Water Works Association (AWWA) publishes a survey on operating data for water utilities in the United States. Using the AWWA survey, revenues from the sales of water to residential and nonresidential users are divided by the amounts of water sold to the respective users. These

¹⁶ Capital cost figures used in the data are, however, historical. Thus, despite the above statement, this study does not claim to have true opportunity costs.

ratios are used as approximate prices for residential and nonresidential customers. The price for residential water usage is $P_R^A = 41.38¢/1000$ gallon and the price for nonresidential water usage is $P_N^A = 20.17¢/1000$ gallon.

The actual price charged to residential users is much higher than the price charged to nonresidential users of water. There may be two reasons for this price difference between residential and nonresidential customers. One possible explanation is the declining block rate schedule adopted by many water utilities. According to the declining block rate, the price is lowered as the amount of consumption of water increases. This is typically done by giving quantity discounts to larger users of water. Although residential users of water are affected by this schedule, their per capita usage is lower than that of industrial and commercial users. Therefore, nonresidential users are in general charged lower prices than residential users of water.

Another possible explanation for the price difference between residential and nonresidential customers is based on the difference in demand elasticities between residential and nonresidential customers. A number of studies have been conducted in the estimation of demand for water; see Mann (1981) for a complete review. In general, price elasticities of both residential and nonresidential demands for water are quite inelastic. The reason is that water expenditures tend to constitute only a small portion of a typical residential user budget and also a relatively minor portion of the operation costs of most commercial and industrial firms. However, the price elasticity of residential demand for water appears to be lower than the price elasticity of

nonresidential demand. The difference in the two elasticities are determined by the essentiality of water and existence of readily accessible substitutes.

Commercial and industrial users of water are more aware of and responsive to changes in price than are residential users. Commercial and industrial firms have possibilities of developing their own supplies of water in the long run in face of a large increase in the price of water and have more substitution possibilities.¹⁷ However, substitution possibilities for residential customers are low, compared with those of nonresidential customers. Therefore, the price elasticity of residential is lower than that of nonresidential water.

The textbook explanation for price difference in two markets having the same cost conditions is based on the difference in price elasticities of demand for residential and nonresidential usage of water. The more elastic a demand, the lower the price in that market. Since the price elasticity of demand for nonresidential water is higher than the price elasticity of demand for residential water, nonresidential customers are charged lower rates than the residential customer.

¹⁷ It is of interest to note a few studies on the industrial demand for water and its substitutability with other inputs, using the translog techniques. Grebenstein and Field (1979) in a study on U.S. manufacturing using cross section data found the own price elasticity of the industrial demand for water to be about -0.5. They also found that water is a complement with capital and labor. Griffin (1977a), however, found that the own price elasticity of the demand to be -2.10 and water is substitute with capital and other inputs in the production of electric power in the U.S.; see also Babin, Willis, and Allen (1982).

Having explained the reasons for the differences in prices of resident and nonresidential water usage,¹⁸ we now turn to compare the actual prices with marginal costs of water for residential and nonresidential customers. To do this, we must examine the data in Table 5.10 on actual price-marginal cost margins and markup ratios for both residential and nonresidential water supply. Price-marginal cost margins and markup ratios provide information about the extent to which the actual prices deviate from marginal costs. Since actual price-marginal cost margins and markup ratios for both residential and nonresidential water supply are greater than zero, this suggests that the actual prices of water supplied to residential and nonresidential customers are in excess of marginal costs. In fact, the residential water supply is priced about 160% above its marginal cost, as evidenced by the markup ratio of residential water supply of 1.58. The nonresidential water supply is priced about 40% above its marginal cost. Further, the markup ratio for residential water supply is higher than that for nonresidential water supply, illustrating the preferential treatment by utilities of nonresidential customers over residential customers.

This information indicates that actual prices depart substantially from the marginal costs of supplying water to consumers. The difficulty of applying marginal cost pricing to the firm experiencing economies of scale is well recognized in the literature, as discussed in Chapter 3.

¹⁸ A third explanation based on the cost allocation method is provided to me by Dick Stevie. If residential users are placing demands at the time of a peak and costs are allocated at the peak, the more cost could be allocated to residential users, hence high price.

Table 5.10

Actual Price-Marginal Cost Margins and Markup

Residential Water Supply:

P_R^A	MC_R	$P_R^A - MC_R$	$\frac{P_R^A - MC_R}{MC_R}$
41.38¢	16.06¢	25.32	1.58

Nonresidential Water Supply:

P_N^A	MC_N	$P_N^A - MC_N$	$\frac{P_N^A - MC_N}{MC_N}$
20.17¢	14.38¢	5.79	0.40

Because of economies of scale, if water utilities were set to their rates to marginal costs, they would earn negative profits (because the revenue yielded from the margin cost pricing falls short of total costs of production of water). The implication is that water utilities have to be subsidized in order to stay in operation. Such subsidization might be possible in the short run. But in the long run, the utility will not be financially viable. Furthermore, the subsidy to water utilities by municipal governments will result in misallocation of resources in other sectors of the economy. Hanke and Davis (1973) argue that the problems in application of marginal cost pricing in the water supply sector involves tradeoffs among price stability, financial viability, economic efficiency, equity, administrative-transaction costs, and political acceptability. The same problems are noted by Mann and Schlenger (1982).

Second Best Pricing Rule

The problem is therefore how to devise a relevant price structure which adequately generates revenue, thereby making the utility financially viable but at the same time achieving economic efficiency. A number of problems with the current declining block rate structure have been noted in the literature. One advantage of using declining block rates is a concern over revenue stability. Water sales have a random element due to both changes in weather and changes in user composition. With the declining block, usage fluctuations occur primarily in the tail blocks thus minimizing revenue instability. The major criticism of declining block rates is their failure to track costs. This means that small users of water tend to subsidize large users. Other problems

related to the declining block include the determination of the number of blocks, consumption break points, and block rate differentials; see Mann (1981).

A water rate design that recently has acquired some support is the inverted block rate, the obverse to the declining block rate. It has generally been advocated as a form of conservation pricing and has been justified on the basis of recent increasing costs associated with capacity expansion. On the premise that water consumption is positively related to income levels, it is also supported as a means of reducing income inequality. However, the inverted rate schedule has the same problem as the declining block rate.

The issue of the second best pricing has been ignored in the water supply literature on the ground that it is difficult to quantify the magnitude of price. In addition, there is the philosophical issue of whether regulators should be concerned with economic inefficiency outside their regulatory jurisdiction (Mann 1981). However, the second best pricing seems to be misunderstood in the water supply literature, obscuring the appropriate calculation. Using the formula (3.89) given in Chapter 3, we can calculate the second best prices for residential and nonresidential usage.

It is appropriate to assume that residential and nonresidential demands for water are not interrelated. Thus, the cross price elasticities of demand for residential and nonresidential water are set to zero in the calculation. The second best prices are evaluated at the sample means of data. Two sets of (own) price elasticities of residential and nonresidential usage are used to see if the second best prices are

sensitive to the variation in the elasticities. There are given in Table 5.11. In case 1, two price elasticities are inelastic with nonresidential price elasticity higher than residential price elasticity. In case 2, nonresidential demand is elastic in relation to the price elasticity of residential water. Even for the difference in the price elasticities of demands in two cases, the second best prices in both cases are not significantly different from each other. Since the price elasticity of nonresidential water is higher in both cases, its percentage deviation from marginal cost is lower than the percentage deviation of the price of residential water from marginal cost. Nevertheless, the second best prices are bounded by the lower limit set by marginal cost and upper limit set by demand elasticities.

Table 5.12 reports the extent to which the actual prices diverge from the second best prices. This table shows that the actual prices are short of the second best prices, but the deviations of the actual prices from the second best prices are small. It suggests, surprisingly, that the actual prices charged to residential and nonresidential consumers are close to the respective second best prices. The upshot is that even if the actual prices of water deviate substantially from marginal costs, they can be considered as efficient prices, constraining the utility to break even. From a review of the literature, it is apparent that no studies have been performed in the area of the second best prices and the relationship to the actual prices of water.

Table 5.11
Second Best Prices Under Alternative Price
Elasticities of Water Demand

Case 1:	$E_{RR} = -0.40$	$E_{NN} = -0.75$
	$P_R = 45.8621¢/1000$	gallons
	$P_N = 22.0097¢/1000$	gallons
Case 2:	$E_{RR} = -0.99$	$E_{NN} = -1.98$
	$P_R = 45.8783¢/1000$	gallons
	$P_N = 21.7799¢/1000$	gallons

Table 5.12

Relation of Actual to Second Best Prices

Residential Water Supply:

P_R^A	P_R^S	$P_R^A - P_R^S$	$\frac{P_R^S - MC_R}{MC_R}$	$\frac{P_R^A - P_R^S}{P_R^S}$
41.38¢	45.86¢	-4.48	1.86	-0.10

Nonresidential Water Supply:

P_N^A	P_N^S	$P_N^A - P_N^S$	$\frac{P_N^S - MC_N}{MC_N}$	$\frac{P_N^A - P_N^S}{P_N^S}$
20.17¢	22.01¢	-1.84	0.53	-0.08

Issue of Cross-Subsidization Among Customer Classes

The extent and form of cross-subsidization is an important concern of utility commissions whose enabling legislation requires that rates be reasonably related to the cost of providing services. It is commonly claimed that residential users of water are subsidizing nonresidential users. One way to examine this claim is to compare total residential revenues with the expenses allocated to the residential water supply. In fact, this is the approach taken in an EPA study (1978). This EPA study, choosing the Cincinnati Water Works as a case study area, concludes that the Cincinnati Water Works runs a deficit on nonresidential water supply and surplus on residential supply. This leads one to conclude that nonresidential customers are subsidized by residential users. This study further points out that "such a result may occur not so much because of cost differences, but because of the rate structure (declining block rates)" and that "for no such subsidization to occur, rates would have to be tailored exactly to costs."

One problem with this approach is that it requires costs to be allocated across outputs. However, because of joint costs present between residential and nonresidential water supply, there is no unambiguous way of correctly separating total costs into residential and nonresidential water supply.¹⁹ The correct segregation of costs would be possible only if costs are nonjoint. Another problem with this approach

¹⁹ For a review on methods of allocation of joint or shared cost in water supply, see EPA study (1978). For an analysis of allocated cost rules and their problems in regulated industries, see Braeutigam (1980).

is that economic analysis of rate structure should begin with marginal costs. Pricing rules grounded in efficiency criteria should be based on marginal costs, thus avoiding any need to separate costs; see Braeutigam (1980), Winston (1981), and Griffin (1982).

We can employ Table 5.10 to investigate the issue of cross subsidization among customer classes. As indicated before, both residential and nonresidential water supply are priced in excess of the respective marginal costs. Further, actual price-marginal cost markup ratios are greater than zero for all customer classes. By implication, there is no evidence of cross subsidization among customer classes in water supply, a result which is contrary to the conclusion of the EPA study (1978). The obvious reason is that the actual prices of residential and nonresidential water supply are tailored proportionally to marginal costs.

Recently, Bruggink (1982) investigated the existence of (third degree) price discrimination among customer classes in municipal utilities. He obtained the same conclusion as drawn above with respect to cross subsidization: no evidence of cross subsidization across the customer classes. However, there are some important differences between his study and this one. As indicated before, his study along with other existing studies are based on an arbitrary functional specification with no theoretical underpinning of the neoclassical cost function. Further, he treats treatment and distribution systems as a separable activity--a restrictive specification. Consequently, the estimates of marginal costs obtained by Bruggink are presumed to be biased.

Is Water Cheap?

A prevailing notion in the water supply field is that water is underpriced or simply that water is too cheap. As a result of the underpricing of water, water utilities have been induced to defer both system maintenance and capital replacement; see Mann and LeFrancois (1982). It is also argued that traditional rate setting methods, used by state regulatory commissions as well as local political bodies, have produced deteriorating water systems characterized by aging capital facilities.

To answer this question, one should first understand what the economic sense is in which water rates can be said to be too low. This study has argued that rates should be based on marginal cost, not average cost. And, actual prices of both residential and nonresidential water usage have been demonstrated to be higher than their respective marginal costs. This implies that water is not underpriced, a result which is contrary to the common argument.

In the comparison of actual prices of water with costs, average cost is, however, commonly employed instead of marginal cost. Thus, it would be useful to obtain an average cost and an "average" marginal cost of water usage and compare them with an average price. Average cost for joint outputs cannot be defined. But we can obtain an approximate average cost without breaking down water into customer classes. An average marginal cost is calculated as a weighted average of marginal costs of residential and nonresidential water usage with weights given by the respective shares of residential and nonresidential water in the

total water produced. It is given by:

$$MC = \frac{Y_R}{Y_R + Y_N} MC_R + \frac{Y_N}{Y_R + Y_N} MC_N$$

The average price for both residential and nonresidential water usage is 30.78¢/1000 gallons. The average cost for both residential and nonresidential water usage at the sample means is $AC = 42.18¢/1000$ gallons. The average marginal cost evaluated at the sample means is $MC = 15.70¢/1000$ gallons. The average cost of water is lower than the average price. This indicates that water is underpriced relative to the average cost of water. Simply put, water is cheap because the cost of service to customers exceeds the actual price charged. In this context, the often heard argument is confirmed.

However, average cost is not the relevant cost concept in the analysis, marginal cost is. The average marginal cost, 15.70¢/1000 gallons, is far less than the actual price and average price. Then, water is overpriced relative to the marginal cost of supply. In other words, the real and opportunity cost of water is much lower than most people in the water supply field expect. This is due to the fact that water utilities experience substantial economies of scale in the provision of water to consumers, as demonstrated before. However, since actual prices charged are higher than the marginal costs, we might say that the saving obtained from economies of scale are not passed to the consuming public in the form of lower prices.

CHAPTER 6

SUMMARY AND CONCLUSIONS

Despite the voluminous literature on other regulated public utilities employing the neoclassical specification of technology, studies on water supply are limited in methodology and scope. These studies have usually consisted of regressions of cost against output, with a number of variations in functional forms and differences in independent variables included. Consequently, the specification of the structure of production is arbitrary and crudely ad hoc with no theoretical reference to the underlying technology.

Previous studies have primarily been concerned with the estimation of scale economies and neglected the role that input prices play in the cost function. Therefore, these studies were not able to address the issue of substitutabilities among the water supply inputs. Also, past studies fail to recognize the fact that the output of a water utility is multidimensional by its nature. In addition, existing studies employ the extremely simplified functional forms such as the Cobb-Douglas. Because of these shortcomings, the previous studies fail to estimate the cost function that appropriately characterize the structure of water supply. As a consequence, the validity of their empirical results remain questionable.

The purpose of this research is to develop a complete economic methodology suitable for applying the neoclassical theory of the multiproduct firm to the analysis to the production structure of water supply. The water supply utility is viewed as a regulated multiproduct firm providing multiple outputs for residential and nonresidential users. Capital, labor, and energy are considered as the major inputs in the production of the delivered water. Particular attention is directed to analyses of production characteristics of the multiproduct firm such as: substitution possibilities among inputs, input requirements, scale economies, and marginal costs. Pricing decisions and cross-subsidization among customer classes are also discussed. In addition, the validity of the restrictive specifications of technology as adopted in the existing literature is examined.

The principle of duality is explicitly utilized to allow the use of the translog multiproduct cost function. An unusual feature of the research conducted with this model is the incorporation of a few operating and spatial characteristics of the firm as measured by capacity utilization and service distance. A cross section of 1973 water utilities in the United States is chosen for this study. In the estimation procedure, the iterative Zellner efficient method is utilized which jointly estimates the translog cost function and the derived input demand functions in order to obtain maximum likelihood estimators.

6.1 MAJOR FINDINGS OF THE RESEARCH

The major findings of the research can be summarized as follows:

1.) Parameter Estimates and Tests of Restrictive Specifications of Technology

Estimation presumes nonhomotheticity among outputs (residential and nonresidential water supply), input prices, and operating variables (capacity utilization and service distance) in the production of delivered water. The translog multiproduct cost function is found to be an adequate description of the technology of water supply. The tests of homogeneity of outputs, input-output separability, nonjointness, and a Cobb-Douglas form produce results that are significantly different from the unconstrained translog cost model. This suggests that existing specifications placing arbitrary restrictions on the translog multiproduct cost function are inappropriate. In addition, the hypothesis that the operating variables should not be included in the model is rejected. Thus, attempts to analyze water supply costs without considering the operating variables would lead to spurious inferences about the structure of water supply.

2.) Input Demand Function for Water Supply

No previous studies have been performed in the estimation of the input demand function for water supply. Capital is found to be a substitute for labor and energy. Substitution between energy and labor is limited to a great extent. Substitution between capital and energy is consistent with the findings of studies in other areas. Energy is an input which requires intensive use in the production process in response to changes in the demand for water and capacity utilization. Further, it is found that capacity utilization and service distance have significant

effects on input requirements. Thus, the inability to account for these variables would lead to a serious problem in the estimation of input requirements for water supply.

3.) Scale Economies

Water utilities are found to experience substantial economies of scale for both residential and nonresidential water supply, a finding which supports previous studies. As conjectured in these studies, the water supply industry is experiencing marked economies associated with the treatment of water but suffers from diseconomies of increasing distribution network size. But the economies in treatment overrides diseconomies in distribution. Further, it is found that scale economies are mainly determined by the provision of water to nonresidential users as opposed to residential users of water.

4.) Marginal Costs and Output Supply

Marginal costs of residential and nonresidential water supply are found to be much lower than the actual prices charged. Marginal costs are quite sensitive to changes in outputs, input prices, and operating variables. Further, from the estimates of marginal costs, the transformation function of water supply is calculated. The transformation function is convex, illustrating decreasing real opportunity costs of residential and nonresidential water supply. This implies that there is a gain from the total specialization in the production of either residential or nonresidential water supply.

5.) Pricing of Water and Cross-Subsidization Among Customer Classes

Since the firm experiences economies of scale, this is suggestive of the infeasibility of marginal cost pricing. The second best pricing rule has been proposed as an alternative scheme through which prices can be calculated for residential and nonresidential water supply. Surprisingly, the existing price structure is close to the second best pricing optimum. Since marginal costs are less than the actual prices, this indicates that real costs of water supply are much lower than what is expected. In addition, this study finds no evidence of cross-subsidization among residential and nonresidential users.

Empirical findings in this research represent first attempts in the water supply area. No comprehensive analysis has been conducted in the past using the econometric methodology adopted in this research. However, before one can draw any sweeping or definite conclusions from these findings, it is important to verify them using different data. This is necessary because of the limitations imposed by the data currently available. Nevertheless, the methodology employed in this research should represent fruitful lines of inquiry into the various issues of water supply.

6.2 SUGGESTIONS FOR FUTURE RESEARCH

The present research uses the cross section samples in order to analyze the long run cost function for water supply. There are, however, a number of potential fruitful lines of extensions.

First, the present research does not assume the technical change in the water supply industry. However, the use of time series data makes possible the analysis of productivity growth and technical progress. Labor productivity has been considered as a sole factor for the overall productivity growth in the water supply literature; see Clark (1979a). But it is more appropriate to estimate total factor or multifactor productivity. Changes in total factor productivity reflect changes in efficiency in the use of all measured inputs rather than in the use of labor inputs alone. A measure of labor productivity can sometimes rise only because capital services or intermediate inputs are being substituted for labor.

Second, this research used the long run cost function to describe the technology of water supply. This implies that firms instantaneously adjust input usages in an optimal fashion. However, estimates of long run cost functions are often subject to bias because firms are not in their long run equilibrium. This suggests that it is desirable to specify and estimate a short run cost function (Eads, Nerlove, and Raduchel 1969). Even if firms are in long run equilibrium, no bias will exist if a short run cost function is specified and estimated. The short run cost function will merely indicate that optimal amounts of the fixed inputs are used. Fortunately, the short run cost function can give the same information concerning technology and input demands as the long run cost function.

Third, in the estimation of the cost function, a major assumption was that the firm seeks to minimize the cost of producing its desired

rate of outputs, subject to the transformation function giving maximum possible outputs for a given set of inputs. Therefore, the firm is assumed to be technically efficient in the sense that it produces maximum outputs for a given input bundle. The firm is also allocatively efficient in the sense that it operates on its least cost expansion path. However, the firm may be inefficient--technically and allocatively--and the observed levels of outputs and inputs may diverge from the optimal levels. Since the estimation of the cost function is carried out with observations on outputs and input prices, one may be estimating the "inefficient" cost function. Recently, a number of studies have been conducted in order to take into account this problem under the estimation of "frontier" production functions; see Aigner, Lovell, and Schmidt (1977), Toda (1976), Schmidt and Lovell (1979), Kopp and Smith (1980), and Greene (1980). Therefore, it seems more appropriate to estimate the cost function using this framework.

Last but not least, the translog is utilized as a functional representation of the cost function in this research. The translog is one of the several flexible functional forms which provide a local approximation to the true underlying cost function. However, the global properties of these flexible forms are usually not known (Caves and Christensen 1980). It is impossible therefore to choose among the available forms on theoretical grounds. It is also impossible to discriminate among them on econometric grounds, since all yield estimation systems which involve similar estimation techniques. Therefore, it is more appropriate to have a method for choosing among these flexible

forms. Recently, some studies have been performed in this area using the Box-Cox transformation function; see Berndt, Diewert and Darrough (1977), Berndt and Khaled (1979), and Appelbaum (1979b). These studies use a generalized Box-Cox functional form for the cost function, which has the translog and other functional forms as special cases. Thus, another appropriate direction in the estimation of the cost function for water supply is to use this technique. While all of the above points may be hard to incorporate in any study, they are nevertheless directions to pursue in future research.

BIBLIOGRAPHY

- Afriat, S.N., "Efficiency Estimation of Production Functions," International Economic Review 13 (3), October 1972.
- Aigner, D.J., C.A.K. Lovell, and P.J. Schmidt, "Formulation and Estimation of Stochastic Frontier Production Function Models," Journal of Econometrics 6 (1), July 1977.
- Allen, R.G.D., Mathematical Analysis for Economists, St. Martin's Press, New York, 1938.
- American Water Works Association, Water Rates Manual.
- Andrews, R., "Economies Associated With Size of Water Utilities and Communities Served in New Hampshire and New England," Water Resources Bulletin, October 1971.
- Appelbaum, E., "Testing Price Taking Behavior," Journal of Econometrics 9, February 1979a.
- Appelbaum, E., "On the Choice of Functional Forms," International Economic Review, June 1979b.
- Appelbaum, E. and U. Kohli, "Canada-United States Trade: Tests For The Small Open-Economy Hypothesis," Canadian Journal of Economics, February 1979.
- Atkinson, S.E. and R. Halvorsen, "A Test of Relative and Absolute Price Efficiency in Regulated Utilities," Review of Economics and Statistics, February 1980.
- Averch, H. and L. Johnson, "Behavior of the Firm Under Regulatory Constraint," American Economic Review, December 1962.
- Babin, F.G., C.E. Willis, and P.G. Allen, "Estimation of Substitution Possibilities Between Water and Other Production Inputs," American Journal of Agricultural Economics, February 1982.
- Bailey, E.E., "Contestability and the Design of Regulatory and Antitrust Policy," American Economic Review, May 1981.
- Bailey, E.E. and A.F. Friedlaender, "Market Structure and Multiproduct Industries," Journal of Economic Literature, September 1982.
- Bard, Y., Nonlinear Parameter Estimation, Academic Press, New York, 1974.
- Barten, A.P., "Maximum Likelihood Estimation of a Complete System of Demand Equations," European Economic Review 1, Fall 1969.

- Baumol, W.J., "Scale Economies, Average Cost and the Profitability of Marginal Cost Pricing," in R.E. Grieson (ed.), Urban Economics and Public Finance: Essays in Honor of William S. Vickrey, Lexington, Mass., D.C. Heath, 1975.
- Baumol, W.J., "On the Proper Cost Tests for Natural Monopoly in a Multiproduct Industry," American Economic Review, December 1977.
- Baumol, W.J., "Contestable Markets: An Uprising in the Theory of Industry," American Economic Review, March 1982.
- Baumol, W.J. and D.F. Bradford, "Optimal Departures From Marginal Cost Pricing," American Economic Review, June 1980.
- Baumol, W.J. and Y.M. Braunstein, "Empirical Study of Scale Economies and Production Complementarity: The Case of Journal Publication," Journal of Political Economy, October 1977.
- Baumol, W.J. and D. Fischer, "Cost-Minimizing Number of Firms and Determination of Industry Structure," Quarterly Journal of Economics, August 1977.
- Baumol, W.J., E.B. Bailey, and R.D. Willig, "Weak Invisible Hand Theorems on the sustainability of Multiproduct Natural Monopoly," American Economic Review, June 1977.
- Baumol, W.J., J.C. Panzar, and R.D. Willig, Contestable Markets and the Theory of Industry Structure, Harcourt Brace Javanovich, New York, 1982.
- Baumol, W.J. and Associates, "The Role of Cost in the Minimum Pricing of Railroad Services," Journal of Business, October 1962.
- Beach, C.M. and J.G. MacKinnon, "Maximum Likelihood Estimation of Singular Equation Systems With Autoregressive Disturbances," International Economic Review, June 1979.
- Beckenstein, A.R., "Scale Economies in the Multiplant Firm: Theory and Empirical Evidence," Bell Journal of Economics, Autumn 1975.
- Berndt, E.R., "Reconciling Alternative Estimates of the Elasticity of Substitution," Review of Economics and Statistics, February 1976.
- Berndt, E.R., "Aggregate Energy, Efficiency, and Productivity Measurement," Annual Review Energy 3, 1978.
- Berndt, E.R. and L.R. Christensen, "The Internal Structure and Functional Relationships: Separability, Substitution, and Aggregation," Review of Economic Studies, July 1973.

- Berndt, E.R. and L.R. Christensen, "The Translog Function and the Substitution of Equipment, Structure, and Labor in U.S. Manufacturing 1929-1968," Journal of Econometrics 1, March 1973.
- Berndt, E.R. and L.R. Christensen, "Testing for the Existence of a Consistent Aggregate Index of Labor Inputs," American Economic Review, June 1974.
- Berndt, E.R. and B.C. Field (eds.), Modeling and Measuring Natural Resource Substitution, MIT Press, Cambridge, MA, 1981.
- Berndt, E.R. and M.S. Khaled, "Parametric Productivity Measurement and Choice Among Flexible Functional Forms," Journal of Political Economy, December 1975.
- Berndt, E.R. and C.J. Morrison, "Income Distribution and Employment Effects of Rising Energy Prices," Resources and Energy 2, 1979.
- Berndt, E.R. and N.E. Savin, "Estimation and Hypothesis Testing in Singular Equation Systems With Autoregressive Disturbances," Econometrica, September 1975.
- Berndt, E.R. and N.E. Savin, "Conflict Among Criteria for Testing Hypotheses in the Multivariate Linear Regression Model," Econometrica, July 1977.
- Berndt, E.R. and D.O. Wood, "Technology, Prices, and the Derived Demand for Energy," Review of Economics and Statistics, August 1975.
- Berndt, E.R. and D.O. Wood, "Engineering and Econometric Interpretation of Energy-Capital Complementarity," American Economic Review, June 1979.
- Berndt, E.R. and D.O. Wood, "Engineering and Econometric Interpretation of Energy-Capital Complementarity: Reply and Further Results," American Economic Review, December 1981.
- Berndt, E.R., M.N. Darrrough, and W.E. Diewert, "Flexible Functional Forms and Expenditure Distributions: An Application to Canadian Consumer Demand Functions," International Economic Review, October 1977.
- Berndt, E.R., M.A. Fuss, and L. Waverman, "Dynamic Models of the Industrial Demand for Energy," Electric Power Research Institute, Palo Alto, California, Technical Report No. EA-580, November 1977.

- Berndt, E.R., C.J. Morrison, and G.C. Watkins, "Dynamic Models of Energy Demand: An Assessment and Comparison," in E.R. Berndt and B.C. Field (eds.), Modeling and Measuring Natural Resource Substitution, MIT Press, Cambridge, MA, 1981.
- Berndt, E.R., B.H. Hall, R.E. Hall, and J.A. Hausman, "Estimation and Inference in Nonlinear Simultaneous Model," Annals of Economic and Social Measurement 3/4, October 1974.
- Binswagen, H.P., "A Cost Function Approach to the Measurement of Elasticities of Factor Demand and Elasticities of Substitution," American Journal of Agricultural Economics 56 (2), May 1974.
- Binswagen, H.P., "The Measurement of Technical Change Biases With Many Factors of Production," American Economic Review, December 1974.
- Blackorby, C. and R.R. Russell, "The Morishima Elasticity of Substitution; Symmetry, Constancy, Separability, and Its Relationship to the Hicks and Allen Elasticities," Review of Economic Studies, January 1981.
- Blackorby, C., D. Primont, and R.R. Russell, "On Testing Separability Restrictions With Flexible Functional Forms," Journal of Econometrics 5 (2), March 1977.
- Blackorby, C.D., D. Primont and R.R. Russell, Duality, Separability and Functional Structure: Theory and Economic Applications, NY, American Elsevier, 1980.
- Braeutigam, R., "An Analysis of Fully Distributed Cost Pricing in Regulated Industries," Bell Journal of Economics, Spring 1980.
- Braeutigam, R.R., A.F. Daugherty, and M.A. Turnquist, "The Estimation of a Hybrid Cost Function for a Railroad Firm," Review of Economics and Statistics, August 1982.
- Brown, L. and G. Chachere, "A Note on the Cost Elasticity-Scale Elasticity Relation," Econometrica, March 1980.
- Brown, R.S., D.W. Caves, and L.R. Christensen, "Modelling the Structure of Cost and Production for Multi-Product Firms," Southern Economic Journal, July 1979.
- Bruggink, T., "Third-Degree Price Discrimination and Regulation in the Municipal Water Industry," Land Economics, February 1982a.
- Bruggink, T.H., "Public Versus Regulated Private Enterprise in the Municipal Water Industry: A Comparison of Operating Costs," Quarterly Review of Economics and Business, 22 (1), Spring 1982b.

- Burgess, D.F., "Duality Theory and Pitfalls in the Specification of Technology," Journal of Econometrics 3 (2), May 1975.
- Burgess, D.F., "A Cost Minimization Approach to Import Demand Functions: Review of Economics and Statistics, March 1974.
- Burgess, D.F., "Tarriffs and Income Distribution: Some Empirical Evidence for the United States," Journal of Political Economy, February 1976.
- Burley, H.T., "Productive Efficiency in U.S. Manufacturing: A Linear Programming Approach," Review of Economics and Statistics, November 1980.
- Caves, D.W. and L.R. Christensen, "The Relative Efficiency of Public and Private Firms in a Competitive Environment: The Case of Canadian Railroads," Journal of Political Economy, October 1980.
- Caves, D.W., L.R. Christensen, and W.E. Diewert, "Multilateral Comparisons of Output, Input and Productivity Using Superlative Numbers," Economic Journal, March 1982.
- Caves, D.W., L.R. Christensen, and J.A. Swanson, "Productivity in U.S. Railroads, 1951-1974," Bell Journal of Economics, Spring 1980.
- Caves, D.W., L.R. Christensen, and J.A. Swanson, "Economic Performance in Regulated and Unregulated Environment: A Comparison of U.S. and Canadian Railroads," Quarterly Journal of Economics, November 1981.
- Caves, D.W., L.R. Christensen, and J.A. Swanson, "Productivity Growth, Scale Economies, and Capacity Utilization in U.S. Railroads, 1955-74," American Economic Review, December 1981.
- Caves, D.W., L.R. Christensen, and M.W. Tretheway, "Flexible Cost Functions for Multiproduct Firms," Review of Economics and Statistics, August 1980.
- Chambers, R.G., "Duality, The Output Effect, and Applied Comparative Statics," American Journal of Agricultural Economics, February 1982.
- Christensen, L.R. and D.W. Caves, "Global Properties of Flexible Functional Forms," American Economic Review, June 1980.
- Christensen, L.R. and W.H. Greene, "Economies of Scale in U.S. Electric Power Generation," Journal of Political Economy, August 1976.

- Christensen, L.R. and W.H. Greene, "An Econometric Assessment of Cost Savings From Coordination in U.S. Electric Power Generation," Land Economics, May 1978.
- Christensen, L.R. and D.W. Jorgenson, "The Measurement of U.S. Real Capital Input, 1927-1967," Review of Income and Wealth 15, December 1969.
- Christensen, L.R. and D.W. Jorgenson, "U.S. Real Product and Real Factor Input, 1929-1927," Review of Income and Wealth, March 1970.
- Christensen, L.R., D.W. Jorgenson, and L.J. Lau, "Conjugate Duality and the Transcendental Logarithmic Function," Econometrica, July 1971.
- Christensen, L.R., D.W. Jorgenson, and L.J. Lau, "Transcendental Logarithmic Production Frontiers," Review of Economics and Statistics, February 1973.
- Clark, R.M., "Water Supply Economics," Journal of the Urban Planning and Development Division, ASCE, August 1976.
- Clark, R.M., "The Safe Drinking Water Act: Its Implications for Planning," in D. Holtz and S. Sebastian (eds.), Municipal Water Systems: The Challenge for Urban Resource Management, Indiana University Press, Bloomington, 1978.
- Clark, R.M., "Labor Wage Rates, Productivity, and the Cost of Water Supply," Journal American Water Works Association 71 (7), July 1979a.
- Clark, R.M., "Water Supply Regionalization: A Critical Evaluation," Journal of the Water Resources Planning and Management Division, ASCE, September 1979b.
- Clark, R.M., "Small Water Systems: Role of Technology," Journal of the Environment Engineering Division, ASCE, 106 (EE1), February 1980.
- Clark, R.M. and R.G. Stevie, "Meeting the Drinking Water Standards: The Price of Regulation," in C.S. Russell (ed.), Safe Drinking Water: Current and Future Problems, Resources for the Future, Washington, DC, 1978.
- Clark, R.M. and R.G. Stevie, "A Water Supply Cost Model Incorporating Spatial Variables," Land Economics, February 1981.
- Clark, R.M., J.I. Gillean, and W.K. Adams, The Cost of Water Supply and Water Utility Management Volume I, EPA-600/5-77-0151, Municipal Environmental Research Laboratory, Office of Research and Development, U.S. Environmental Protection Agency, Cincinnati, Ohio, November 1977.

- Clark, R.M., H. Youn Kim, and G.S. Huband, "Energy Utilization in Water Treatment and Delivery," in "Energy Conservation in the Design of Water Quality Control Facilities," Proceeding of Fourth Mid-America Conference on Environmental Design, held by the Department of Civil Engineering, University of Missouri-Columbia, Columbia, MO 65211, 1979.
- Clark, R.M., R. Stevie, and G. Trygg, "The Cost of Municipal Water Supply: A Case Study," EPA-600/2/-76-179, Municipal Environmental Research Laboratory, Office of Research and Development, U.S. Environmental Protection Agency, Cincinnati, Ohio, July 1976.
- Coase, R.H., "The Economics of Uniform Pricing Systems," Manchester School of Economics and Social Studies, May 1947.
- Corbo, V. and P. Meller, "The Translog Production Function," Journal of Econometrics 10, June 1979.
- Cornell, Holly, "A Consultant Looks at Future Water Utility Energy Problems: Journal American Water Works Association, April 1978.
- Courville, L., "Regulation and Efficiency in the Electric Utility Industry," Bell Journal of Economics and Management Science, Spring 1974.
- Cowing, T.G. and R.E. Stevenson (eds.), Productivity Measurement in Regulated Industries, Academic Press, New York, 1981.
- Dajani, J.S., "Cost Studies of Urban Public Services," Land Economics, November 1973.
- Dajani, J.S. and R.S. Gemmel, "Economic Guidelines for Public Utilities Planning," Journal of the Urban Planning and Development Division, ASCE, September 1976.
- Darrough, M. and J.M. Heineke, "The Multi-Output Translog Production Cost Function: The Case of Law Enforcement Agencies," in J.M. Heineke (ed.), Economic Models of Criminal Behavior, North Holland, Amsterdam, 1978.
- DeVany, A.S., "Capacity Utilization Under Alternative Regulatory Constraints: An Analysis of Taxi Markets," Journal of Political Economy, February 1975.
- Denny, M., "The Relationship Between Functional Forms for the Production System," Canadian Journal of Economics, February 1974.
- Denny, M. and D. May, "The Existence of a Real Value-Added Function in the Canadian Manufacturing Sector," Journal of Econometrics 5 (1), January 1977.

- Denny, M. and C. Pinto, "An Aggregate Model With Multiple-Product Technologies," in M.A. Fuss and D. McFadden (eds.), Production Economics: A Dual Approach to Theory and Applications, North Holland, Amsterdam, 1978.
- Denny, M., C. Everson, M. Fuss, and L. Waverman, "Estimating the Effects of Diffusion of Technological Innovations in Telecommunications: The Production Structure of Bell Canada," Canadian Journal of Economics, February 1981.
- Diewert, W.E., "An Application of the Shephard Theorem: A Generalized Leontief Production Function," Journal of Political Economy, May 1971.
- Diewert, W.E., "Separability and a Generalization of the Cobb-Douglas Cost Production, and Indirect Utility Function," Technical Report 86, IMSSS, Stanford University, January 1973a.
- Diewert, W.E., "Functional Forms for Profits and Transformation Functions," Journal of Economic Theory 6, June 1973b.
- Diewert, W.E., "Applications of Duality Theory," in M.D. Intriligator and D.A. Kendrick (eds.), Frontiers of Quantitative Economics Volume II, Amsterdam, North Holland, 1974a.
- Diewert, W.E., "A Note on Aggregation and Elasticities of Substitution," Canadian Journal of Economics, February 1974b.
- Diewert, W.E., "Exact and Superlative Index Numbers," Journal of Econometrics, May 1976.
- Diewert, W.R., "Superlative Index Numbers and Consistency in Aggregation," Econometrica 46 (4), July 1978a.
- Diewert, W.R., "Duality Approaches to Microeconomic Theory," IMSSS Technical Report 281, Stanford University, 1978b.
- Diewert, W.E., "Capital and the Theory of Productivity Measurement," American Economic Review, May 1980.
- Diewert, O.W. and C. Parkan, "Linear Programming Tests of Regularity Conditions for Production Functions," Discussion paper No. 79-01, University of British Columbia, Vancouver, Canada, January 1979.
- Eads, G., M. Nerlove, and W. Raduchel, "A Long-Run Cost Function for the Local Service Airline Industry: An Experiment in Non-Linear Estimation," Review of Economics and Statistics, August 1969.

- Everson, C.K., "Dynamic Demand for Utilization, Maintenance Installations and Retirements of Railroad Freight Cars," International Economic Review, June 1982.
- Faulhaber, G.R., "Cross-Subsidization: Pricing in Public Enterprises," American Economic Review, December 1975.
- Feldman, S.L., "On the Peak Load Pricing of Urban Water Supply," Water Resources Research, April 1975.
- Feldstein, M.S., "Distributional Equity and the Optimal Structure of Public Prices," American Economic Review, March 1982.
- Field, B.C. and C. Grebenstein, "Capital-Energy Substitution in U.S. Manufacturing," Review of Economics and Statistics, August 1980.
- Ford, J.L. and J.J. Warford, "Cost Functions for the Water Industry," Journal of Industrial Economics, November 1969.
- Friedlander, A.F. and R.H. Spady, "A Derived Demand Function For Freight Transportation," Review of Economics and Statistics, August 1980.
- Fuss, M.A., "The Demand for Energy in Canadian Manufacturing," Journal of Econometrics 5 (1), January 1977.
- Fuss, M.A. "The Derived Demand for Energy in the Presence of Supply Constraints," in W.T. Ziemba, S.L. Schwartz, and E. Koenigsberg (eds.), Energy Policy Modelling: United States and Canadian Experiences, Vol I, Martinus Nijhoff, Boston, 1978.
- Fuss, M.A. and D. McFadden (eds.), Production Economics: A Dual Approach to Theory and Applications, North Holland, Amsterdam, 1978.
- Gamponia, V. and G. Brown, "Steel and Energy Substitution in U.S. Manufacturing," Southern Economic Journal, January 1982.
- Gollop, F.M. and D.W. Jorgenson, "U.S. Productivity Growth by Industry," in J.W. Kendrick and B.N. Vaccara (eds.), New Developments in Productivity Measurement and Analysis, National Bureau of Economic Research, University of Chicago Press, Chicago, 1980.
- Gollop, F.M. and S.H. Karlson, "An Intertemporal Model of Behavior: The Regulated Monopolist," SSRI Workshop Paper 7812, University of Wisconsin, Madison, Wisconsin.
- Gollop, F.M. and S.H. Karlson, "The Impact of the Fuel Adjustment Mechanism on Economic Efficiency," Review of Economics and Statistics 60, November 1978.

- Grebenstein, C.R. and B.C. Field, "Substituting for Water Inputs in U.S. Manufacturing," Water Resources Research, April 1979.
- Greene, W.H., "On the Estimation of a Flexible Frontier Production Model," Journal of Econometrics 13, 1980, 101-115.
- Griffin, J.M., "The Effects of Higher Prices on Electric Consumption," Bell Journal of Economics and Management, Autumn 1974.
- Griffin, J.M., "Statistical Cost Analysis," Quarterly Journal of Economics, February 1977a.
- Griffin, J.M., "Long-Run Production Modelling with Pseudo Data: Electric Power Generation," Bell Journal of Economics, Spring 1977b.
- Griffin, J.M., "Inter-Fuel Substitution Possibilities: A Translog Application to Intercountry Data," International Economic Review, October 1977c.
- Griffin, J.M., "The Econometrics of Joint Production: Another Approach," Review of Economics and Statistics, November 1977d.
- Griffin, J.M., "Joint Production Technology: The Case of Petrochemicals," Econometrica, March 1978.
- Griffin, J.M., "The Energy-Capital Complementary Controversy: A Progress Report on Reconciliation Attempts," in E.R. Berndt and B.C. Field (eds.), Modeling and Measuring Natural Resource Substitution, MIT Press, Cambridge, MA, 1981a.
- Griffin, J.M., "Engineering and Econometric Interpretations of Energy - Capital Complementarity: Comment," American Economic Review, December 1981b.
- Griffin, J.M., "The Welfare Implications of Externalities and Price Elasticities for Telecommunications Pricing," Review of Economics and Statistics, February 1982.
- Griffin, J.M. and P.R. Gregory, "An Intercountry Translog Model of Energy Substitution Responses," American Economic Review, December 1976.
- Guilkey, D.K. and C.A.K. Lovell, "On the Flexibility of the Translog Approximation," International Economic Review, February 1979.
- Hall, R.E., "The Specification of Technology With Several Kinds of Output," Journal of Political Economy, August 1973.
- Halvorsen, R., "Energy Substitution in U.S. Manufacturing," Review of Economics and Statistics, November 1977.

- Halvorsen, R. and J. Ford, "Substitution Among Energy, Capital and Labor Inputs in U.S. Manufacturing," in R. Pindyck (ed.), Advances in the Economics of Energy and Resources, Vol. 1, Greenwich, Connecticut, J.A.I. Press, 1979.
- Harmatuck, D.J., "A Motor Carrier Joint Cost Function," Journal of Transport Economics and Policy, May 1981.
- Hanke, S.H., "Pricing Urban Water," in S. Mushkin (ed.), Public Prices for Public Products, The Urban Institute, Washington, DC, 1972.
- Hanke, S.H., "Water Rates: An Assessment of Current Issues," Journal American Water Works Association, May 1975.
- Hanke, S.H., "A Method for Integrating Engineering and Economic Planning," Journal American Water Works Association, September 1978.
- Hanke, S.H. and R.K. Davis, "Demand Management Through Responsive Pricing," Journal American Water Works Association, September 1971.
- Hanke, S.H. and R.K. Davis, "Potential for Marginal Cost Pricing in Water Resource Management," Water Resources Research, August 1973.
- Hanoch, G., "Homotheticity in Joint Production," Journal of Economic Theory 2, 1970.
- Hanoch, G., "Production and Demand Models With Direct or Indirect Implicit Additivity," Econometrica, May 1975a.
- Hanoch, G., "The Elasticity of Scale and the Shape of the Average Costs," American Economic Review 65, December 1975b.
- Hanoch, G. and M. Rothschild, "Testing the Assumptions of Production Theory: A Nonparametric Approach," Journal of Political Economy 80, March/April 1972.
- Hasenkamp, G., "A Study of Multiple-Output Production Functions: Klein's Railroad Study Revisited," Journal of Econometrics 4 (3), August 1973.
- Helms, B.P. and R.M. Clark, "Financing Municipal Water Supply," Journal American Water Works Association, May 1978.
- Hines, L.G., "The Long-Run Cost Function of Water Production for Selected Wisconsin Communities," Land Economics, February 1969.
- Hinomoto, H., "Unit and Total Cost Functions for Water Treatment Based on Keonig's Data," Water Resources Research, October 1971.

- Hirshleifer, J., J.D. De Haven, and J.W. Milliman, Water Supply: Economics, Technology, and Policy, Chicago, IL, University of Chicago Press, 1960.
- Howe, C.W., "Water Pricing in Residential Areas," Journal American Water Works Association, May 1968.
- Hudson, E.A. and D.W. Jorgenson, "U.S. Energy Policy and Economic Growth, 1975-2000," Bell Journal of Economics and Management, Autumn 1974.
- Hulten, C.R., "Divisia Index Numbers," Econometrica 41, November 1973.
- Hulten, C.R., "Technical Change and the Reproducibility of Capital," American Economic Review 65, December 1975.
- Hulten, C.R., "On the 'Importance' of Productivity Change," American Economic Review 69, March 1979.
- Humphrey, D.B., "Economics of Scale in Federal Reserve Check Processing Operations," Journal of Econometrics, January 1981.
- Humphrey, A.B. and J.R. Moroney, "Substitution Among Capital, Labor and Natural Resource Products in American Manufacturing," Journal of Political Economy, February 1975.
- Jorgenson, D.W., "Capital Theory and Investment Behavior," American Economic Review, May 1963.
- Journal American Water Works Association, Special Issues on "Regionalization of Water Utilities," December 1979, and Regionalization: Why and How," May 1981.
- Kahn, A., "Applications of Economics to an Imperfect World," American Economic Review, May 1979.
- Kako, T., "An Application of the Decomposition Analysis of Derived Demand for Factor Inputs in U.S. Manufacturing," Review of Economics and Statistics, May 1980.
- Kelley, C.W., "Pricing of Water," Journal American Water Works Association, January 1977.
- Kendrick, J.W. and B.N. Vaccara (eds.), New Developments in Productivity Measurement and Analysis, National Bureau of Economic Research, University of Chicago Press, Chicago, 1980.
- Kesselman, J., S.H. Williamson, and E.R. Berndt, "Tax Credit for Employment Rather than Investment," American Economic Review, June 1977.

- Kang, H. and G.M. Brown, "Partial and Full Elasticities of Substitution and the Energy-Capital Complementarity Controversy," in E.R. Berndt and B.C. Field (eds.), Modeling and Measuring Natural Resource Substitution, MIT Press, Cambridge, MA, 1981.
- Kim, H.Y., "Energy Substitution, Economies of Scale, and Optimal Planning in U.S. Water Utilities," a research proposal for a Cooperative Agreement submitted to U.S. Environmental Protection Agency, June 1980.
- Kmenta, J. and R. Gilbert, "Small Sample Properties of Alternative Estimators of Seemingly Unrelated Regressions," Journal of the American Statistical Association 63, December 1968.
- Kohli, U.R., "A Gross National Product Function and the Derived Demand or Imports and Support of Exports," Canadian Journal of Economics, May 1978.
- Kohli, U.R., "Nonjointness and Factor Intensity in U.S. Production," International Economic Review, February 1981.
- Konig, L., "Cost of Water Treatment by Coagulation, Sedimentation, and Rapid-Sand Filtration," Journal American Water Works Association, March 1967.
- Kopp, R.J. and V.K. Smith, "Frontier Production Function Estimates for Steam Electric Generation: A Comparative Analysis," Southern Economic Journal, April 1980.
- Lau, L.J., "Profit Functions of Technologies With Multiple Inputs and Outputs," Review of Economic and Statistics 54 (3), 1972.
- Lau, L.J., "Comments on Applications of Duality Theory," in M.D. Intriligator and D.A. Kendrick (eds.), Frontiers of Quantitative Economics, Vol II, North Holland, Amsterdam, 1974.
- Lau, L.J., "A Characterization of the Normalized Restricted Profit Functions," Journal of Economic Theory 12, 1976.
- Lau, L.J., "Applications of Profit Functions," in M. Fuss and D. McFadden (eds.), Production Economics: A Dual Approach to Theory and Applications, North Holland, Amsterdam, 1978a.
- Lau, L.J., "Testing and Imposing Monotonicity, Convexity, and Quasi-Convexity Constraints," in M. Fuss and D. McFadden (eds.), Production Economics: A Dual Approach to Theory and Applications, North Holland, Amsterdam, 1978.
- Lau, L.J., "On Exact Index Numbers," Review of Economics and Statistics, February 1979.

- Levy, V., "Total Factor Productivity, Non-Neutral Technical Change and Economic Growth," Journal of Development Economics 8 (1), February 1981.
- Linaweaver, F. P. and C.J. Clark, "Costs of Water Transmission," Journal American Water Works Association, December 1964.
- Lopez, R.E. "The Structure of Production and the Derived Demand for Inputs in Canadian Agriculture," American Journal of Agricultural Economics, February 1980.
- Magnus, J.R., "Substitution Between Energy and Non-Energy Inputs in the Netherlands 1950-1976," International Economic Review, June 1979.
- Malinvaud, E., Statistical Methods of Econometrics, North Holland, Amsterdam, 1980.
- Mann, P.C. "Water Service: Regulation and Rate Reform," Occasional Paper No. 4, The National Regulatory Research Institute, OSU, November 1981.
- Mann, P.C. and P.R. LeFrancois, "The Real Price of Urban Water," Journal American Water Works Association, January 1982.
- Mann, P.C. and J.L. Mikesell, Ownership and Water System Operation," Water Resources Bulletin, October 1976.
- Mann, P.C. and J.L. Mikesell, "Cost Behavior and Water Systems," Water Resources Bulletin, October 1979.
- Mann, P.C. and D.L. Schlenger, "Marginal Cost and Seasonal Pricing of Water Service," Journal American Water Works Association, January 1982.
- May, J.D. and M. Denny, "Factor-Augmenting Technical Progress and Productivity in U.S. Manufacturing," International Economic Review, October 1979.
- McFadden, D., "Production Functions with Constant Elasticities of Substitution," Review of Economic Studies 30, 1963.
- McFadden, D., "Cost, Revenue, and Profit Functions," in M. Fuss and D. McFadden (eds.), Production Economics: A Dual Approach to Theory and Applications, North Holland, Amsterdam, 1978.
- Marguardt, W., "A Critique of Some Ridge Regression Methods: Comment," Journal of the American Statistical Association, March 1980.
- Marquardt, D.W. and R.D. Snee, "Ridge Regression in Practice," The American Statistician 29, 1975.

- Moroney, J.R. and A.L. Toevs, "Factor Costs and Factor Use: An Analysis of Labor, Capital, and Natural Resource Inputs," Southern Economic Journal, October 1977.
- Moroney, J.R. and A.L. Toevs, "Input Prices, Substitution, and Product Inflation," in R. Pindyck (ed.), Advances in the Economics of Energy and Resources, Vol 1, Greenwich, Connecticut, J.A.I. Press, 1979.
- Moroney, J.R. and J.M. Trapani, "Alternative Models of Substitution and Technical Change in Natural Resource - Intensive Industries," in E. Berndt and B. Field (eds.), Modeling and Measuring Natural Resource Substitution, Cambridge, MIT Press, 1981.
- Moroney, J.R. and J.M. Trapani, "Factor Demand and Substitution in Mineral-Intensive Industries," Bell Journal of Economics, Spring 1981.
- Mountain, D.C., "The Spatial Distribution of Electricity Demand: Its Impact Upon Input Usage," Land Economics, February 1981.
- Mountain, D.C., "Imported Electricity: A Substitute for Inputs Used by the Regional Electricity Industry," Journal of Regional Science, February 1982.
- Mundlak, Y., "Elasticity of Substitution and the Theory of Derived Demand," Review of Economic Studies 35 (2), April 1968.
- Murray, M.P., R. Spann, L. Pulley, and E. Beauvais, "The Demand for Electricity in Virginia," Review of Economics and Statistics, November 1978.
- Nadiri, M.I., "Some Approaches to the Theory of Total Factor Productivity: A Survey," Journal of Economic Literature, December 1970.
- Nerlove, M., "Returns to Scale in Electricity Supply," in C.F. Christ (ed.), Measurement in Economics - Studies in Mathematical Economics and Econometrics in Memory of Yehuda Grunfeld, Stanford, Stanford University Press, 1963.
- Neuberg, L.G., "Two Issues in the Municipal Ownership of Electric Power Distribution Systems," Bell Journal of Economics, Spring 1977.
- Oberhofer, W. and J. Kmenta, "A General Procedure for Obtaining Maximum Likelihood Estimates in Generalized Regression Models," Econometrica, May 1974.
- Ohta, M., "Production Technologies of the U.S. Boiler and Turbogenerator Industries and Hedonic Price Indexes for Their Products: A Cost-Function Approach," Journal of Political Economy, February 1975.

- Ordover, J.A. and R.D. Willig, "On the Optimal Production of Journals and Sometimes Shares Goods," American Economic Review, June 1978.
- Orlob, G.T. and M.R. Lindork, "Cost of Water Treatment in California," Journal American Water Works Association, January 1958.
- Oum, Tae Hoon, "A Cross Sectional Study of Freight Transportation Demand and Rail-Truck Competition in Canada," Bell Journal of Economics, Autumn 1979.
- Paul, E.S., "Pricing Rules and Efficiency," in S. Mushkin (ed.), Public Prices for Public Products, The Urban Institute, Washington, DC, 1972.
- Panzar, J.C. and R.D. Willig, "Free Entry and the Sustainability of Natural Monopoly," Bell Journal of Economics, Spring 1977a.
- Panzar, J.C. and R.D. Willig, "Economies of Scale in Multi-Output Production," Quarterly Journal of Economics, August 1977b.
- Panzar, J.C. and R.D. Willig, "Theoretical Determinants of the Industrial Demand for Electricity by Time of Day," Journal of Econometrics, January 1979.
- Panzar, J.C. and R.D. Willig, "Economies of Scope," American Economic Review, May 1981.
- Parks, R.W., "Price Responsiveness of Factor Utilization in Swedish Manufacturing, 1850-1950," Review of Economics and Statistics, May 1971.
- Petersen, H.D., "An Empirical Test of Regulatory Effects," Bell Journal of Economics, Spring 1975.
- Phillips, S.L., "Factor Demand in the Provision of Public Safety," in J.M. Heineke (ed.), Economic Models of Criminal Behavior, North Holland, Amsterdam, 1978.
- Pindyck, R.S., "Interfuel Substitution and the Industrial Demand for Energy: An International Comparison," Review of Economics and Statistics, May 1979.
- Pindyck, R.S., The Structure of World Energy Demand, MIT Press, Cambridge, Mass., 1980.
- Pittman, R.W., "Issue in Pollution Control: Interplant Cost Differences and Economies of Scale," Land Economics, February 1981.

- Powell, A.A. and F.H.G. Gruen, "The Constant Elasticity of Transformation Production Frontier and Linear Supply System," International Economic Review, October 1980.
- Ray, S.C., "A Translog Cost Function Analysis of U.S. Agriculture, 1938-77," American Journal of Agricultural Economics, August 1982.
- Ramsey, F., "A Contribution to the Theory of Taxation," Economic Journal, March 1927.
- Reiter, G., "Survey and Comparison of California Water Agencies Rate Structure," Journal American Water Works Association, July 1977.
- Russell, R., "Functional Separability and Partial Elasticities of Substitution," Review of Economic Studies 42 (1), 1975.
- Sakai, Y., "Substitution and Expansion Effects in Production Theory: The Case of Joint Production," Journal of Economic theory 9, 1974, 255-274.
- Samuelson, P.A., "The Fundamental Singularity Theorem of Nonjoint Production," International Economic Review 7 (1), 1966.
- Samuelson, P.A. and S. Swamy, "Invariant Economic Index Numbers and Canonical Duality: Survey and Synthesis," American Economic Review, September 1974.
- Sato, K. and T. Koizumi, "On the Elasticities of Substitution and Complementarity," Oxford Economic Paper, March 1973.
- Schmalensee, Richard, "A Note on Economies of Scale and Natural Monopoly in the Distribution of Public Utility Services," Bell Journal of Economics, Spring 1978.
- Schmidt, P., "On the Statistical Estimation of Parametric Frontier Production Functions," Review of Economics and Statistics, May 1976.
- Schmidt, P. and C.A.K. Lovell, "Estimating Technical and Allocative Inefficiency Relative to Stochastic Production and Cost Frontiers," Journal of Econometrics 9 (3), February 1979.
- Shephard, R.W., Cost and Production Functions, Princeton University Press, Princeton, New Jersey, 1953.
- Shephard, R.W., Theory of Cost and Production Functions, Princeton University Press, Princeton, New Jersey, 1970.
- Smith, V.K., "Elasticities of Substitution for a Regulated Cost Function," Economic Letters 7, 1981, 215-219.

- Spady, R.H. and A.F. Friedlaender, "Hedonic Cost Function for the Regulated Trucking Industry," Bell Journal of Economics, Spring 1973.
- Spann, Robert M., "Rate of Return Regulation and Efficiency in Production: An Empirical Test of the Averch-Johnson Thesis," Bell Journal of Economics and Management Science, Spring 1974.
- Stapleton, D.C., "Inferring Long-Term Substitution Possibilities from Cross-Section and Time-Series Data," in E.R. Berndt and B.C. Field (eds.), Modeling and Measuring Natural Resource Substitution, MIT Press, MA, 1981.
- Star, S. and R.E. Hall, "An Approximate Divisia Index of Total Factor Productivity," Econometrica, March 1976.
- Stevenson, R.E., "Measuring Technological Bias," American Economic Review, March 1980.
- Stevie, R.G. and R.M. Clark, "Cost for Small Systems to Meet the National Interim Drinking Water Regulation," Journal American Water Works Association, January 1982.
- Stevie, R.G. R.M. Clark, and J.Q. Adams, Managing Small Water Systems: A Cost Study, Volume I, EPA-600/2-79-147a, Municipal Environmental Research Laboratory, Office of Research and Development, U.S. Environmental Protection Agency, Cincinnati, Ohio, September 1979.
- Teece, D.J., "Economies of Scope and the Scope of the Enterprise," Journal of Economic Behavior and Organization, September 1980.
- Theil, H., Principles of Econometrics, John Wiley, New York, 1971.
- Theil, H., "A New Index Number Formula," Review of Economics and Statistics, November 1973.
- Toda, Y., "Estimation of a Cost Function When the Cost is not Minimum: The Case of Soviet Manufacturing Industries, 1958-1971," Review of Economics and Statistics, August 1976.
- Turvey, R., "Analyzing the Marginal Cost of Water Supply," Land Economics, May 1976.
- U.S. Environmental Protection Agency, Survey of Operating and Financial Characteristics of Community Water Systems, report prepared for the Environment Protection Agency by Temple, Barker, and Sloane, Inc., Office of Water Supply, Washington, DC, April 1977.

- U.S. Environmental Protection Agency, Planning Water Supply: Cost-Rate Differentials and Plumbing Permits by H.C. Goddard, R.G. Stevie, and G.D. Trygg, EPA-600/5-78-008, MERL, Office of Research and Development, U.S. Environmental Protection Agency, Cincinnati, Ohio, May 1978.
- Uri, N.D., "Regional Interfuel Substitution by Electric Companies: The Short-Term Prospects," Annals of Regional Science, July 1978.
- Uzawa, H., "Production Functions with Constant Elasticities of Substitution," Review of Economic Studies 29 (4), October 1962.
- Uzawa, H., "Duality Principles in the Theory of Cost and Production," International Economic Review 5 (2), October 1964.
- Vickrey, W., "Some Implications of Marginal Cost Pricing for Public Utilities," American Economic Review, May 1955.
- Vickrey, W., "Economic Efficiency and Pricing," in S. Mushkin (ed.), Public Prices for Public Products, The Urban Institute, Washington, DC, 1972.
- Wales, T.J., "On the Flexibility of Flexible Functional Forms: An Empirical Approach," Journal of Econometrics 5 (2), March 1977.
- Willig, R.D., "Multiproduct Technology and Market Structure," American Economic Review, May 1979.
- Willig, R.D. and E.E. Bailey, "The Economic Gradient Method," American Economic Review, May 1979.
- Wills, J., "Technical Change in the U.S. Primary Metals Industry," Journal of Econometrics 10, 1979.
- Winston, C., "The Welfare Effects of ICC Rate Regulation Revisited," Bell Journal of Economics, Spring 1981.
- Woodland, A.D., "Substitution of Structures, Equipment and Labor in Canadian Production," International Economic Review, February 1975.
- Woodland, A.D., "Modelling the Production Sector of an Economy: A Selective Survey and Analysis," Discussion paper No. 76-21, Department of Economics, University of British Columbia, May 1976.
- Woodland, A.D., "A Dual Approach to Equilibrium in the Production Sector in International Trade Theory," Canadian Journal of Economics, February 1977.

Woodland, A.D., "Stochastic Specification and the Estimation of Share Equations," Journal of Econometrics, August 1979.

Zellner, A., "An Efficient Method of Estimating Seemingly Unrelated Regression and Tests Aggregation Bias," Journal of the American Statistical Association, June 1962.

Zellner, A., "Estimates for Seemingly Unrelated Regression Equations: Some Exact Finite Sample Results," Journal of the American Statistical Association, December 1973.

Ziemba, W.T., S.L. Schwartz, and E. Koenigsberg (eds.), "Energy Policy Modeling: United States and Canadian Experience, Vol. I, Martin Nijhoff, Boston, 1978.